

```

***** SAS program
*****;

/** For Table 3 & Table 5 */

/** The line size is needed to ensure it is wide enough for the output to look
good */
options nocenter nodate ls=120 ps=80 nonumber;
/* read in main and validation data */
libname j "/udd/strui/sofie/N02";

/** read in validation study which should include the true exposure, surrogate
and the same covariates that
appear in the primary regression model */

data valid;
    set j.expono2;
run;

/** For Table 3 of the paper */

title "The multivariate measurement error model for the true exposure in
relation to the surrogate";
title2 "===== For table 3
=====";
proc reg data=valid;
    model trueno2=surrno2 smokers gascook unvented sex age pet;
run;

/** read in the main study which should include the outcome, the surrogate and
covariates. Note that
    we have one data set for each outcome */
/** main_conjWheeze is for current conjunctivitis and current wheeze;
    main_phlegm is for current phlegm
    main_ige is for total IgE;
*/

data main_conjWheeze;
    set j.epi2;
run;

data main_phlegm;
    set j.epiphlegm2;
run;

data main_ige;
    set j.epiige2;
run;

/** Output from the primary log-binomial regression is input to the measurement
error
    correction macro.

    You can use the older version 'make' statement to output parameter estimates
and
    variance-covariance matrix, which will not be supported by future SAS
versions.

    You can use ods instead. Syntax for IgE as the outcome:
ods output ParameterEstimates=b_ige;

```

```

ods output covb=cov_ige;
**/

title "===== For table 5
=====";
title2 'For Wheeze, proc genmod, log binomial regression unadjusted for
measurement error';
proc genmod data=main_conjWheeze descending;
class RESPNR;
model wheeze = surrno2 smokers gascook unvented sex age pet/ dist=binomial
link=log covb;
make 'ParameterEstimates' out=b_Wheeze;
make 'covb' out=cov_Wheeze;
run;

title2 'For Conjunctivitis and Wheeze, proc genmod, log binomial regression
unadjusted for measurement error';
proc genmod data=main_conjWheeze descending;
class RESPNR;
model conj= surrno2 smokers gascook unvented sex age pet/ dist=binomial link=log
covb;
make 'ParameterEstimates' out=b_conj;
make 'covb' out=cov_conj;
run;

title2 'For Current Phlegm, proc genmod, log binomial unadjusted for measurement
error';
proc genmod data=main_phlegm descending;
class RESPNR;
model phlegm = surrno2 smokers gascook unvented sex age pet/ dist=binomial
link=log covb;
make 'ParameterEstimates' out=b_phlegm;
make 'covb' out=cov_phlegm;
run;

title2 'For total IgE, proc genmod, log binomial unadjusted for measurement
error';
proc genmod data=main_ige descending;
class RESPNR;
model ige = surrno2 smokers gascook unvented sex age pet/ dist=binomial link=log
covb;
make 'ParameterEstimates' out=b_ige;
make 'covb' out=cov_ige;
run;

/** increments for uncorrected analysis. The order should be exactly same as the
previous
unadjusted proc genmod model with intercept added as the first line, that
is,
intercept, surrogate and the covariates */

data inc_uncor;
input name $ incrementt;
datalines;
intercept 1
surrno2 17.6
smokers 1
gascook 1
unvented 1
sex 1

```

```

    age 1
    pet 1
    ;
run;

/** increments for the corrected analysis. The order is same as the above by
replacing
    surrogate with true exposure, that is, intercept, true exposure, and the
covariates **/
data inc_cor;
    input name $ increment;
    cards;
        intercept 1
    trueno2 17.6
    smokers 1
    gascook 1
    unvented 1
    sex 1
    age 1
    pet 1
    ;
run;

/** call blinplus8 macro for each health outcome **/
%include "/udd/strui/sofie/blinplus8.sas";

title "For Current Wheeze";
%blinplus8( type = GENMOD,
            valid = valid,
            main_est = b_wheeze,
            main_var = cov_wheeze,
            increments1 = inc_uncor,
            increments2 = inc_cor,
            err_var = surrno2 smokers gascook unvented sex age pet,
            true_var = trueno2 smokers gascook unvented sex age pet,
            depend = wheeze
        );

title "For Current Conjunctivitis";
%blinplus8( type = GENMOD,
            valid = valid,
            main_est = b_conj,
            main_var = cov_conj,
            increments1 = inc_uncor,
            increments2 = inc_cor,
            err_var = surrno2 smokers gascook unvented sex age pet,
            true_var = trueno2 smokers gascook unvented sex age pet,
            depend = conj
        );

title "For Current Phlegm";
%blinplus8( type = GENMOD,
            valid = valid,
            main_est = b_phlegm,
            main_var = cov_phlegm,
            increments1 = inc_uncor,
            increments2 = inc_cor,

```

```

err_var = surrno2 smokers gascook unvented sex age pet,
true_var = trueno2 smokers gascook unvented sex age pet,
depend  = phlegm
);

```

```

title "For total IgE";
%blinplus8( type      = GENMOD,
            valid     = valid,
            main_est  = b_ige,
            main_var  = cov_ige,
            increments1 = inc_uncor,
            increments2 = inc_cor,
            err_var   = surrno2 smokers gascook unvented sex age pet,
            true_var  = trueno2 smokers gascook unvented sex age pet,
            depend    = ige
);

```

```
***** SAS lst *****
*****;
```

The multivariate measurement error model for the true exposure in relation to the surrogate

===== For table 3
=====

The REG Procedure

Model: MODEL1

Dependent Variable: TRUENO2 trueno2

Number of Observations Read 67
Number of Observations Used 67

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	3402.41084	486.05869	12.34	<.0001
Error	59	2324.49497	39.39822		
Corrected Total	66	5726.90581			

Root MSE	6.27680	R-Square	0.5941
Dependent Mean	23.68709	Adj R-Sq	0.5460
Coeff Var	26.49883		

Parameter Estimates

Variable Pr > t	Label	DF	Parameter Estimate	Standard Error	t Value
Intercept 0.4594	Intercept	1	6.66672	8.95294	0.74
SURRNO2 0.0004	surrno2	1	0.42166	0.11342	3.72
SMOKERS 0.2561	smokers	1	1.93825	1.69026	1.15
GASCOOK 0.1755	gascook	1	2.53185	1.84651	1.37
UNVENTED <.0001	unvented	1	24.89096	3.17355	7.84
SEX 0.1364	geslacht 0=v 1=m	1	-2.41091	1.59667	-1.51
AGE 0.8785	age	1	-0.13263	0.86365	-0.15
PET 0.2427	pet	1	-1.92010	1.62716	-1.18

===== For table 5

For Wheeze, proc genmod, log binomial regression unadjusted for measurement error

The GENMOD Procedure

Model Information

Data Set	WORK.MAIN_CONJWHEEZE
Distribution	Binomial
Link Function	Log
Dependent Variable	WHEEZE wheeze

Number of Observations Read	1936
Number of Observations Used	1927
Number of Events	182
Number of Trials	1927
Missing Values	9

Class Level Information

Class	Levels	Values
RESPNR	1936	21401 21402 21403 21404 21405 21406 21407 21408 21409 21411 21501 21502 21503 21504 21505 21506 21507 21508 21509 21510 21511 21512 21513 21601 21603 21604 21605 21701 21702 21703 21705 21707 21802 21803 21804 21805 21806 21807 21808 21810 21811 21812 ...

Response Profile

Ordered Value	WHEEZE	Total Frequency
1	1	182
2	0	1745

PROC GENMOD is modeling the probability that WHEEZE='1'.

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	SURRNO2
Prm3	SMOKERS
Prm4	GASCOOK
Prm5	UNVENTED
Prm6	SEX
Prm7	AGE
Prm8	PET

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	1919	1192.7645	0.6216
Scaled Deviance	1919	1192.7645	0.6216
Pearson Chi-Square	1919	1928.8453	1.0051
Scaled Pearson X2	1919	1928.8453	1.0051
Log Likelihood		-596.3822	

Algorithm converged.

===== For table 5

For Wheeze, proc genmod, log binomial regression unadjusted for measurement error

The GENMOD Procedure

Estimated Covariance Matrix						
Prm6	Prm1 Prm7	Prm2 Prm8	Prm3	Prm4	Prm5	
Prm1	0.45074	-0.006419	-0.000174	-0.02066	0.0002446	
-0.01288	-0.02036	-0.003611				
Prm2	-0.006419	0.0001825	-0.000175	-0.000243	9.5274E-6	
0.0000225	0.0000338	-0.000066				
Prm3	-0.000174	-0.000175	0.02119	0.0001008	-0.003294	-
0.000148	-0.000409	-0.003141				
Prm4	-0.02066	-0.000243	0.0001008	0.03896	-0.004831	-
0.000681	-0.000305	-0.000014				
Prm5	0.0002446	9.5274E-6	-0.003294	-0.004831	0.05041	
0.0001236	-0.000112	0.001359				
Prm6	-0.01288	0.0000225	-0.000148	-0.000681	0.0001236	
0.02023	0.0000576	0.001337				
Prm7	-0.02036	0.0000338	-0.000409	-0.000305	-0.000112	
0.0000576	0.002093	-0.000400				
Prm8	-0.003611	-0.000066	-0.003141	-0.000014	0.001359	
0.001337	-0.000400	0.02049				

Analysis Of Parameter Estimates							
Parameter > ChiSq	DF	Estimate	Standard Error	Wald	95% Confidence Limits	Chi- Square	Pr
Intercept	1	-3.4997	0.6714	-4.8156	-2.1839	27.17	
<.0001							
SURRN02	1	0.0258	0.0135	-0.0007	0.0523	3.65	
0.0560							
SMOKERS	1	0.0008	0.1456	-0.2845	0.2861	0.00	
0.9957							
GASC00K	1	-0.0104	0.1974	-0.3973	0.3764	0.00	
0.9579							
UNVENTED	1	0.2388	0.2245	-0.2013	0.6789	1.13	
0.2876							
SEX	1	0.2731	0.1422	-0.0057	0.5519	3.69	
0.0549							
AGE	1	0.0250	0.0458	-0.0647	0.1147	0.30	
0.5846							
PET	1	-0.2733	0.1432	-0.5538	0.0073	3.64	
0.0563							
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

===== For table 5

For Conjunctivitis and Wheeze, proc genmod, log binomial regression unadjusted
for measurement error

The GENMOD Procedure

Model Information

Data Set	WORK.MAIN_CONJWHEEZE
Distribution	Binomial
Link Function	Log
Dependent Variable	CONJ conj

Number of Observations Read	1936
Number of Observations Used	1936
Number of Events	142
Number of Trials	1936

Class Level Information

Class	Levels	Values
RESPNR	1936	21401 21402 21403 21404 21405 21406 21407 21408 21409 21411 21501 21502 21503 21504 21505 21506 21507 21508 21509 21510 21511 21512 21513 21601 21603 21604 21605 21701 21702 21703 21705 21707 21802 21803 21804 21805 21806 21807 21808 21810 21811 21812 ...

Response Profile

Ordered Value	CONJ	Total Frequency
1	1	142
2	0	1794

PROC GENMOD is modeling the probability that CONJ='1'.

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	SURRN02
Prm3	SMOKERS
Prm4	GASCOOK
Prm5	UNVENTED
Prm6	SEX
Prm7	AGE
Prm8	PET

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
-----------	----	-------	----------

Deviance	1928	994.6144	0.5159
Scaled Deviance	1928	994.6144	0.5159
Pearson Chi-Square	1928	1921.8347	0.9968
Scaled Pearson X2	1928	1921.8347	0.9968
Log Likelihood		-497.3072	

Algorithm converged.

===== For table 5

For Conjunctivitis and Wheeze, proc genmod, log binomial regression unadjusted
for measurement error

The GENMOD Procedure

Estimated Covariance Matrix						
Prm6	Prm1 Prm7	Prm2 Prm8	Prm3	Prm4	Prm5	
Prm1	0.61721	-0.008392	-0.000669	-0.03470	0.0002427	
-0.01789	-0.02748	-0.003161				
Prm2	-0.008392	0.0002357	-0.000199	-0.000402	-3.075E-7	
0.0000402	0.0000501	-0.000149				
Prm3	-0.000669	-0.000199	0.02792	0.0004790	-0.004576	-
0.000349	-0.000652	-0.004234				
Prm4	-0.03470	-0.000402	0.0004790	0.06369	-0.006290	-
0.000653	-0.000682	0.0000159				
Prm5	0.0002427	-3.075E-7	-0.004576	-0.006290	0.10176	-
0.000659	-0.000010	0.002711				
Prm6	-0.01789	0.0000402	-0.000349	-0.000653	-0.000659	
0.02651	0.0000970	0.002252				
Prm7	-0.02748	0.0000501	-0.000652	-0.000682	-0.000010	
0.0000970	0.002765	-0.000542				
Prm8	-0.003161	-0.000149	-0.004234	0.0000159	0.002711	
0.002252	-0.000542	0.02714				

Analysis Of Parameter Estimates							
Parameter > ChiSq	DF	Estimate	Standard Error	Wald	95% Confidence Limits	Chi- Square	Pr
Intercept	1	-5.5986	0.7856	-7.1384	-4.0588	50.78	
<.0001							
SURRN02	1	0.0452	0.0154	0.0151	0.0753	8.67	
0.0032							
SMOKERS	1	0.0324	0.1671	-0.2951	0.3599	0.04	
0.8462							
GASC00K	1	0.2609	0.2524	-0.2337	0.7556	1.07	
0.3012							
UNVENTED	1	-0.3152	0.3190	-0.9404	0.3100	0.98	
0.3231							
SEX	1	0.2906	0.1628	-0.0285	0.6098	3.19	
0.0743							
AGE	1	0.1211	0.0526	0.0181	0.2242	5.30	
0.0213							
PET	1	-0.2303	0.1647	-0.5532	0.0926	1.95	
0.1621							
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

===== For table 5

=====

For Current Phlegm, proc genmod, log binomial unadjusted for measurement error

The GENMOD Procedure

Model Information

Data Set	WORK.MAIN_PHLEGM
Distribution	Binomial
Link Function	Log
Dependent Variable	phlegm phlegm

Number of Observations Read	1907
Number of Observations Used	1907
Number of Events	180
Number of Trials	1907

Class Level Information

Class	Levels	Values
RESPNR	1907	21401 21402 21403 21404 21405 21406 21407 21408 21409 21411 21501 21502 21503 21504 21505 21506 21507 21508 21509 21510 21511 21512 21513 21601 21603 21604 21605 21701 21702 21703 21705 21707 21802 21803 21804 21805 21806 21807 21808 21810 21811 21812 ...

Response Profile

Ordered Value	phlegm	Total Frequency
1	1	180
2	0	1727

PROC GENMOD is modeling the probability that phlegm='1'.

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	surrno2
Prm3	smokers
Prm4	gascook
Prm5	unvented
Prm6	SEX
Prm7	age
Prm8	pet

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
-----------	----	-------	----------

Deviance	1899	1175.7322	0.6191
Scaled Deviance	1899	1175.7322	0.6191
Pearson Chi-Square	1899	1897.7155	0.9993
Scaled Pearson X2	1899	1897.7155	0.9993
Log Likelihood		-587.8661	

Algorithm converged.

===== For table 5

For Current Phlegm, proc genmod, log binomial unadjusted for measurement error

The GENMOD Procedure

Estimated Covariance Matrix

Prm6	Prm1 Prm7	Prm2 Prm8	Prm3	Prm4	Prm5	
Prm1	0.46084	-0.006465	-0.001527	-0.02537	0.002120	
-0.01166	-0.02069	-0.002500				
Prm2	-0.006465	0.0001832	-0.000143	-0.000267	-0.000010	
0.0000252	0.0000321	-0.000057				
Prm3	-0.001527	-0.000143	0.02117	0.0001231	-0.002801	-
0.000536	-0.000371	-0.003249				
Prm4	-0.02537	-0.000267	0.0001231	0.04649	-0.004445	-
0.000953	-0.000533	0.0000111				
Prm5	0.002120	-0.000010	-0.002801	-0.004445	0.06649	-
0.000900	-0.000188	0.001013				
Prm6	-0.01166	0.0000252	-0.000536	-0.000953	-0.000900	
0.02014	0.0000683	0.001532				
Prm7	-0.02069	0.0000321	-0.000371	-0.000533	-0.000188	
0.0000683	0.002152	-0.000476				
Prm8	-0.002500	-0.000057	-0.003249	0.0000111	0.001013	
0.001532	-0.000476	0.02072				

Analysis Of Parameter Estimates

Parameter > ChiSq	DF	Estimate	Standard Error	Wald	95% Confidence Limits	Chi- Square	Pr
Intercept	1	-3.8023	0.6789	-5.1329	-2.4718	31.37	
<.0001							
surrno2	1	0.0321	0.0135	0.0056	0.0586	5.63	
0.0177							
smokers	1	0.0123	0.1455	-0.2729	0.2975	0.01	
0.9327							
gascook	1	0.2173	0.2156	-0.2053	0.6399	1.02	
0.3135							
unvented	1	-0.1080	0.2579	-0.6133	0.3974	0.18	
0.6755							
SEX	1	0.0995	0.1419	-0.1786	0.3776	0.49	
0.4832							
age	1	0.0336	0.0464	-0.0573	0.1245	0.53	
0.4684							
pet	1	-0.4248	0.1439	-0.7069	-0.1427	8.71	
0.0032							
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

===== For table 5

=====

For total IgE, proc genmod, log binomial unadjusted for measurement error

The GENMOD Procedure

Model Information

Data Set	WORK.MAIN_IGE
Distribution	Binomial
Link Function	Log
Dependent Variable	IgE IgE

Number of Observations Read	817
Number of Observations Used	817
Number of Events	327
Number of Trials	817

Class Level Information

Class	Levels	Values
RESPNR	817	21501 21502 21503 21504 21507 21512 21603 21605 21701 21707 21804 21805 21807 21810 21811 21812 21814 21822 22502 22504 22505 22506 22507 22508 22510 22512 22513 22515 22529 22530 22531 22602 22603 22604 22606 22607 22608 22613 22615 22616 22618 22619 ...

Response Profile

Ordered Value	Ig E	Total Frequency
1	1	327
2	0	490

PROC GENMOD is modeling the probability that IgE='1'.

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	surrno2
Prm3	smokers
Prm4	gascook
Prm5	unvented
Prm6	SEX
Prm7	age
Prm8	pet

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
-----------	----	-------	----------

Deviance	809	1062.0029	1.3127
Scaled Deviance	809	1062.0029	1.3127
Pearson Chi-Square	809	819.8081	1.0134
Scaled Pearson X2	809	819.8081	1.0134
Log Likelihood		-531.0015	

Algorithm converged.

===== For table 5

For total IgE, proc genmod, log binomial unadjusted for measurement error

The GENMOD Procedure

Estimated Covariance Matrix

Prm6	Prm1 Prm7	Prm2 Prm8	Prm3	Prm4	Prm5	
Prm1 0.0003010	0.20847 -0.01085	-0.002434 0.0001137	-0.000834	-0.01015	-0.001657	
Prm2 0.000066	-0.002434 0.0000202	0.0000656 -0.000091	-0.000014	-0.000054	0.0001303	-
Prm3 0.000024	-0.000834 -0.000258	-0.000014 -0.001350	0.007960	-0.000433	-0.001423	-
Prm4 0.000352	-0.01015 0.0000382	-0.000054 -0.000089	-0.000433	0.01437	-0.001060	-
Prm5 0.000566	-0.001657 -0.000344	0.0001303 0.001059	-0.001423	-0.001060	0.01472	-
Prm6 0.007190	0.0003010 -0.000213	-0.000066 0.0003566	-0.000024	-0.000352	-0.000566	
Prm7 0.000213	-0.01085 0.001033	0.0000202 -0.000031	-0.000258	0.0000382	-0.000344	-
Prm8 0.0003566	0.0001137 -0.000031	-0.000091 0.007222	-0.001350	-0.000089	0.001059	

Analysis Of Parameter Estimates

Parameter > ChiSq	DF	Estimate	Standard Error	Wald	95% Confidence Limits	Chi- Square	Pr
Intercept <.0001	1	-2.3229	0.4566	-3.2178	-1.4280	25.88	
surrno2 <.0001	1	0.0344	0.0081	0.0185	0.0503	18.02	
smokers 0.0884	1	0.1520	0.0892	-0.0228	0.3269	2.90	
gascook 0.9192	1	0.0122	0.1199	-0.2228	0.2471	0.01	
unvented 0.0063	1	0.3314	0.1213	0.0935	0.5692	7.46	
SEX 0.0067	1	0.2297	0.0848	0.0635	0.3959	7.34	
age 0.8039	1	0.0080	0.0321	-0.0550	0.0710	0.06	
pet 0.0139	1	-0.2090	0.0850	-0.3756	-0.0424	6.05	
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

For Current Wheeze

MEASUREMENT ERROR CORRECTION OF REGRESSION ESTIMATES VIA THE METHOD
OF REGRESSION CALIBRATION

References: Rosner BA, Spiegelman D, Willett WC. Amer J Epi
1990. 132:734-745.

Spiegelman D, Kipnis V, Carroll R.

unpublished manuscript

Programmers: Donna Spiegelman, Sc.D., Harvard School of
Public Health, Boston MA

Doug Grove, M.S., Aidan McDermott, M.A.,
Carrie Wager, B.S.,

of The Channing Laboratory, Boston MA.

TYPE OF REGRESSION: GENMIX

Dependent Variable: wheeze

OPTIONS USED:

Validation Study Regression Coefficients:

	intercep	surrno2	smokers	gascook	unvented	sex	age	pet
surrno2	6.666717	0.421658	1.938248	2.531852	24.89096	-2.41091	-0.13263	-1.9201

Variance matrix of error-variables:

surrno2

surrno2 39.39822

Main study regression coefficients: Uncorrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00000	1.00	-3.49972	0.67137	0.03021	0.00810 - 0.11261	
surrno2 0.05604	17.60	0.02581	0.01351	1.57508	0.98835 - 2.51011	
smokers 0.99571	1.00	0.00078	0.14557	1.00078	0.75236 - 1.33123	
gascook 0.95793	1.00	-0.01041	0.19738	0.98964	0.67215 - 1.45710	
unvented 0.28757	1.00	0.23878	0.22453	1.26970	0.81767 - 1.97163	
sex 0.05490	1.00	0.27306	0.14224	1.31398	0.99428 - 1.73648	
age 0.58464	1.00	0.02501	0.04575	1.02533	0.93738 - 1.12152	
pet 0.05629	1.00	-0.27326	0.14316	0.76090	0.57473 - 1.00736	

Main study regression coefficients: Corrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00014	1.00	-3.90784	1.02398	0.02008	0.00270 - 0.14944	
trueno2 0.08924	17.60	0.06122	0.03602	2.93711	0.84772 - 10.17626	
smokers 0.54656	1.00	-0.11787	0.19550	0.88881	0.60589 - 1.30384	
gascook 0.51175	1.00	-0.16541	0.25210	0.84755	0.51710 - 1.38917	
unvented 0.17461	1.00	-1.28498	0.94656	0.27666	0.04327 - 1.76876	
sex 0.02894	1.00	0.42065	0.19258	1.52295	1.04414 - 2.22132	
age 0.63596	1.00	0.03313	0.06999	1.03368	0.90118 - 1.18567	
pet 0.39610	1.00	-0.15571	0.18349	0.85580	0.59729 - 1.22621	

For Current Conjunctivitis

MEASUREMENT ERROR CORRECTION OF REGRESSION ESTIMATES VIA THE METHOD OF REGRESSION CALIBRATION

References: Rosner BA, Spiegelman D, Willett WC. Amer J Epi 1990. 132:734-745.

Spiegelman D, Kipnis V, Carroll R.

unpublished manuscript

Programmers: Donna Spiegelman, Sc.D., Harvard School of Public Health, Boston MA

Doug Grove, M.S., Aidan McDermott, M.A.,
Carrie Wager, B.S.,
of The Channing Laboratory, Boston MA.

TYPE OF REGRESSION: GENMIX

Dependent Variable: conj

OPTIONS USED:

Validation Study Regression Coefficients:

	intercep	surrno2	smokers	gascook	unvented	sex	age	pet
surrno2	6.666717	0.421658	1.938248	2.531852	24.89096	-2.41091	-0.13263	-1.9201

Variance matrix of error-variables:

surrno2

surrno2 39.39822

Main study regression coefficients: Uncorrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00000	1.00	-5.59858	0.78563	0.00370	0.00079 - 0.01727	
surrno2 0.00323	17.60	0.04521	0.01535	2.21610	1.30489 - 3.76358	
smokers 0.84617	1.00	0.03241	0.16708	1.03294	0.74449 - 1.43317	
gascook 0.30119	1.00	0.26092	0.25237	1.29812	0.79158 - 2.12880	
unvented 0.32307	1.00	-0.31522	0.31900	0.72963	0.39045 - 1.36344	
sex 0.07427	1.00	0.29065	0.16283	1.33729	0.97190 - 1.84007	
age 0.02126	1.00	0.12111	0.05258	1.12875	1.01821 - 1.25130	
pet 0.16215	1.00	-0.23030	0.16475	0.79430	0.57511 - 1.09703	

Main study regression coefficients: Corrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00001	1.00	-6.31342	1.41874	0.00181	0.00011 - 0.02923	
trueno2 0.02098	17.60	0.10723	0.04645	6.60076	1.32953 - 32.77107	
smokers 0.50794	1.00	-0.17542	0.26496	0.83911	0.49920 - 1.41045	
gascook 0.97601	1.00	-0.01056	0.35119	0.98949	0.49714 - 1.96947	
unvented 0.01722	1.00	-2.98419	1.25288	0.05058	0.00434 - 0.58946	
sex 0.03426	1.00	0.54916	0.25941	1.73180	1.04157 - 2.87944	
age 0.20307	1.00	0.13534	0.10632	1.14492	0.92955 - 1.41020	
pet 0.92148	1.00	-0.02441	0.24765	0.97588	0.60061 - 1.58564	

For Current Phlegm

MEASUREMENT ERROR CORRECTION OF REGRESSION ESTIMATES VIA THE METHOD
OF REGRESSION CALIBRATION

References: Rosner BA, Spiegelman D, Willett WC. Amer J Epi
1990. 132:734-745.

Spiegelman D, Kipnis V, Carroll R.

unpublished manuscript

Programmers: Donna Spiegelman, Sc.D., Harvard School of
Public Health, Boston MA

Doug Grove, M.S., Aidan McDermott, M.A.,
Carrie Wager, B.S.,
of The Channing Laboratory, Boston MA.

TYPE OF REGRESSION: GENMIX

Dependent Variable: phlegm

OPTIONS USED:

Validation Study Regression Coefficients:

	intercep	surrno2	smokers	gascook	unvented	sex	age	pet
surrno2	6.666717	0.421658	1.938248	2.531852	24.89096	-2.41091	-0.13263	-1.9201

Variance matrix of error-variables:
surrno2

surrno2 39.39822

Main study regression coefficients: Uncorrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00000	1.00	-3.80232	0.67886	0.02232	0.00590 - 0.08444	
surrno2 0.01769	17.60	0.03211	0.01354	1.75957	1.10313 - 2.80664	
smokers 0.93272	1.00	0.01228	0.14551	1.01236	0.76115 - 1.34648	
gascook 0.31348	1.00	0.21732	0.21561	1.24274	0.81443 - 1.89632	
unvented 0.67546	1.00	-0.10795	0.25785	0.89767	0.54154 - 1.48801	
sex 0.48322	1.00	0.09950	0.14191	1.10462	0.83640 - 1.45885	
age 0.46843	1.00	0.03363	0.04639	1.03420	0.94432 - 1.13264	
pet 0.00317	1.00	-0.42479	0.14394	0.65390	0.49316 - 0.86704	

Main study regression coefficients: Corrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00012	1.00	-4.30995	1.11862	0.01343	0.00150 - 0.12034	
trueno2 0.04554	17.60	0.07614	0.03808	3.81947	1.02694 - 14.20566	
smokers 0.52011	1.00	-0.13530	0.21036	0.87345	0.57833 - 1.31918	
gascook 0.93072	1.00	0.02454	0.28223	1.02484	0.58941 - 1.78197	
unvented 0.04863	1.00	-2.00322	1.01593	0.13490	0.01842 - 0.98807	
sex 0.17153	1.00	0.28307	0.20703	1.32720	0.88453 - 1.99142	
age 0.58663	1.00	0.04373	0.08043	1.04470	0.89234 - 1.22307	
pet 0.16077	1.00	-0.27859	0.19864	0.75685	0.51277 - 1.11711	

For total IgE

MEASUREMENT ERROR CORRECTION OF REGRESSION ESTIMATES VIA THE METHOD
OF REGRESSION CALIBRATION

References: Rosner BA, Spiegelman D, Willett WC. Amer J Epi
1990. 132:734-745.

Spiegelman D, Kipnis V, Carroll R.

unpublished manuscript

Programmers: Donna Spiegelman, Sc.D., Harvard School of
Public Health, Boston MA

Doug Grove, M.S., Aidan McDermott, M.A.,
Carrie Wager, B.S.,
of The Channing Laboratory, Boston MA.

TYPE OF REGRESSION: GENMIX

Dependent Variable: ige

OPTIONS USED:

Validation Study Regression Coefficients:

	intercep	surrno2	smokers	gascook	unvented	sex	age	pet
surrno2	6.666717	0.421658	1.938248	2.531852	24.89096	-2.41091	-0.13263	-1.9201

Variance matrix of error-variables:

surrno2

surrno2 39.39822

Main study regression coefficients: Uncorrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00000	1.00	-2.32291	0.45658	0.09799	0.04004 - 0.23978	
surrno2 0.00002	17.60	0.03438	0.00810	1.83146	1.38503 - 2.42178	
smokers 0.08837	1.00	0.15203	0.08922	1.16420	0.97742 - 1.38667	
gascook 0.91923	1.00	0.01215	0.11987	1.01223	0.80029 - 1.28030	
unvented 0.00632	1.00	0.33136	0.12133	1.39285	1.09806 - 1.76680	
sex 0.00675	1.00	0.22970	0.08479	1.25823	1.06557 - 1.48571	
age 0.80387	1.00	0.00798	0.03214	1.00801	0.94647 - 1.07356	
pet 0.01393	1.00	-0.20899	0.08499	0.81141	0.68691 - 0.95847	

Main study regression coefficients: Corrected

	INCREMENT	B	SE	RR	95% CI	p
intercep 0.00282	1.00	-2.86651	0.95986	0.05690	0.00867 - 0.37337	
trueno2 0.00516	17.60	0.08154	0.02915	4.19999	1.53629 - 11.48213	
smokers 0.97233	1.00	-0.00601	0.17318	0.99401	0.70791 - 1.39574	
gascook 0.35523	1.00	-0.19429	0.21016	0.82342	0.54542 - 1.24311	
unvented 0.02858	1.00	-1.69822	0.77574	0.18301	0.04001 - 0.83713	
sex 0.01006	1.00	0.42629	0.16563	1.53156	1.10699 - 2.11896	
age 0.80762	1.00	0.01880	0.07719	1.01897	0.87590 - 1.18542	
pet 0.74260	1.00	-0.05242	0.15963	0.94893	0.69399 - 1.29751	

***** SAS log *****

*****;

1 The SAS System

10:51 Tuesday, October 2, 2007

NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.

NOTE: SAS (r) 9.1 (TS1M3)

Licensed to HARVARD UNIV - SCHOOL OF PUBLIC HEALTH - T&R, Site 0001177014.

NOTE: This session is executing on the SunOS 5.9 platform.

NOTE: (E9BX01) SAS 9.1.3 SP 4

You are running SAS 9. Some SAS 8 files will be automatically converted by the V9 engine; others are incompatible. Please see <http://support.sas.com/rnd/migration/planning/platform/64bit.html>

PROC MIGRATE will preserve current SAS file attributes and is recommended for converting all your SAS libraries from any SAS 8 release to SAS 9. For details and examples, please see <http://support.sas.com/rnd/migration/index.html>

Channing Laboratory

NOTE: SAS initialization used:

real time 3.19 seconds

cpu time 0.68 seconds

1

2 /** For Table 3 & Table 5 **/

3

4 /** The line size is needed to ensure it is wide enough for the
eoutput to look good **/

5 options nocenter nodate ls=120 ps=80 nonumber;

6 /* read in main and validation data */

7 libname j "/udd/strui/sofie/N02";

NOTE: Libref J was successfully assigned as follows:

Engine: V9

Physical Name: /udd/strui/sofie/N02

8

9 /** read in validation study which should include the true exposure,
surrogate and the same covariates that

10 appear in the primary regression model **/

11

12 data valid;

13 set j.expono2;

NOTE: Data file J.EXPON02.DATA is in a format native to another host or the file encoding does not match the session

encoding. Cross Environment Data Access will be used, which may require additional CPU resources and reduce performance.

14 run;

NOTE: There were 67 observations read from the data set J.EXPON02.

NOTE: The data set WORK.VALID has 67 observations and 17 variables.

NOTE: DATA statement used (Total process time):

real time	1.00 seconds
cpu time	0.17 seconds

```
15
16      /** For Table 3 of the paper**/
17
18      title "The multivariate measurement error model for the true exposure
in relation to the surrogate";
19      title2 "===== For table 3
=====";
20      proc reg data=valid;
21          model trueno2=surrno2 smokers gascook unvented sex age pet;
22      run;

23
24      /** read in the main study which should include the outcome, the
surrogate and covariates. Note that
25          we have one data set for each outcome **/
26      /** main_conjWheeze is for current conjunctivitis and current wheeze;
27          main_phlegm is for current phlegm
28          main_ige is for total IgE;
29      **/
30
```

NOTE: The PROCEDURE REG printed page 1.

NOTE: PROCEDURE REG used (Total process time):

real time	1.79 seconds
cpu time	0.40 seconds

The SAS System

```
31      data main_conjWheeze;  
32      set j.epi2;
```

NOTE: Data file J.EPI2.DATA is in a format native to another host or the file encoding does not match the session encoding. Cross Environment Data Access will be used, which may require additional CPU resources and reduce performance.

```
33      run;
```

NOTE: There were 1936 observations read from the data set J.EPI2.

NOTE: The data set WORK.MAIN_CONJWHEEZE has 1936 observations and 26 variables.

NOTE: DATA statement used (Total process time):

real time	0.50 seconds
cpu time	0.16 seconds

```
34  
35      data main_phlegm;  
36      set j.epiphlegm2;
```

NOTE: Data file J.EPIPHLEGM2.DATA is in a format native to another host or the file encoding does not match the session encoding. Cross Environment Data Access will be used, which may require additional CPU resources and reduce performance.

```
37      run;
```

NOTE: There were 1907 observations read from the data set J.EPIPHLEGM2.

NOTE: The data set WORK.MAIN_PHLEGM has 1907 observations and 26 variables.

NOTE: DATA statement used (Total process time):

real time	0.48 seconds
cpu time	0.13 seconds

```
38  
39      data main_ige;  
40      set j.epiige2;
```

NOTE: Data file J.EPIIGE2.DATA is in a format native to another host or the file encoding does not match the session encoding. Cross Environment Data Access will be used, which may require additional CPU resources and reduce performance.

```
41      run;
```

NOTE: There were 817 observations read from the data set J.EPIIGE2.

NOTE: The data set WORK.MAIN_IGE has 817 observations and 26 variables.

NOTE: DATA statement used (Total process time):

real time	0.27 seconds
cpu time	0.11 seconds

```
42  
43      /** Output from the primary log-binomial regression is input to the  
44      measurement error  
45      correction macro.  
46      You can use the older version 'make' statement to output  
parameter estimates and
```

```

47          variance-covariance matrix, which will not be supported by
future SAS versions.
48
49          You can use ods instead. Syntax for IgE as the outcome:
50          ods output ParameterEstimates=b_ige;
51          ods output covb=cov_ige;
52      **/
53
54      title "===== For table 5
=====";
55      title2 'For Wheeze, proc genmod, log binomial regression unadjusted
for measurement error';
56      proc genmod data=main_conjWheeze descending;
57      class RESPNR;
58      model wheeze = surrno2 smokers gascook unvented sex age pet/
dist=binomial link=log covb;
59      make 'ParameterEstimates' out=b_Wheeze;
60      make 'covb' out=cov_Wheeze;
61      run;

```

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: Class levels for some variables were not printed due to excessive size.

NOTE: PROC GENMOD is modeling the probability that WHEEZE='1'.

NOTE: Algorithm converged.

NOTE: The scale parameter was held fixed.

NOTE: The data set WORK.B_WHEEZE has 9 observations and 8 variables.

NOTE: The data set WORK.COV_WHEEZE has 8 observations and 9 variables.

NOTE: The PROCEDURE GENMOD printed pages 2-3.

NOTE: PROCEDURE GENMOD used (Total process time):

The SAS System

real time	2.65 seconds
cpu time	1.39 seconds

```
62
63      title2 'For Conjunctivitis and Wheeze, proc genmod, log binomial
regression unadjusted for measurement error'
64      ! ;
65      proc genmod data=main_conjWheeze descending;
66      class RESPNR;
67      model conj= surrno2 smokers gascook unvented sex age pet/
dist=binomial link=log covb;
68      make 'ParameterEstimates' out=b_conj;
69      make 'covb' out=cov_conj;
70      run;
```

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: Class levels for some variables were not printed due to excessive size.

NOTE: PROC GENMOD is modeling the probability that CONJ='1'.

NOTE: Algorithm converged.

NOTE: The scale parameter was held fixed.

NOTE: The data set WORK.B_CONJ has 9 observations and 8 variables.

NOTE: The data set WORK.COV_CONJ has 8 observations and 9 variables.

NOTE: The PROCEDURE GENMOD printed pages 4-5.

NOTE: PROCEDURE GENMOD used (Total process time):

real time	1.76 seconds
cpu time	1.17 seconds

```
70
71      title2 'For Current Phlegm, proc genmod, log binomial unadjusted for
measurement error';
72      proc genmod data=main_phlegm descending;
73      class RESPNR;
74      model phlegm = surrno2 smokers gascook unvented sex age pet/
dist=binomial link=log covb;
75      make 'ParameterEstimates' out=b_phlegm;
76      make 'covb' out=cov_phlegm;
77      run;
```

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: Class levels for some variables were not printed due to excessive size.

NOTE: PROC GENMOD is modeling the probability that phlegm='1'.

NOTE: Algorithm converged.

NOTE: The scale parameter was held fixed.

NOTE: The data set WORK.B_PHLEGM has 9 observations and 8 variables.

NOTE: The data set WORK.COV_PHLEGM has 8 observations and 9 variables.

NOTE: The PROCEDURE GENMOD printed pages 6-7.

NOTE: PROCEDURE GENMOD used (Total process time):

real time	2.48 seconds
cpu time	1.16 seconds

```

78
79         title2 'For total IgE, proc genmod, log binomial unadjusted for
measurement error';
80         proc genmod data=main_ige descending;
81         class RESPNR;
82         model ige = surrno2 smokers gascook unvented sex age pet/
dist=binomial link=log covb;
83         make 'ParameterEstimates' out=b_ige;
84         make 'covb' out=cov_ige;
85         run;

```

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: The MAKE statement is obsolete and may not be supported in future releases. Use the ODS OUTPUT statement instead.

NOTE: Class levels for some variables were not printed due to excessive size.

NOTE: PROC GENMOD is modeling the probability that IgE='1'.

NOTE: Algorithm converged.

NOTE: The scale parameter was held fixed.

NOTE: The data set WORK.B_IGE has 9 observations and 8 variables.

NOTE: The data set WORK.COV_IGE has 8 observations and 9 variables.

NOTE: The PROCEDURE GENMOD printed pages 8-9.

NOTE: PROCEDURE GENMOD used (Total process time):

real time	0.68 seconds
-----------	--------------

cpu time	0.52 seconds
----------	--------------

The SAS System

```
87
88      /** increments for uncorrected analysis. The order should be exactly
89      same as the previous
90      unadjusted proc genmod model with intercept added as the first
91      line, that is,
92      intercept, surrogate and the covariates **/
93
94      data inc_uncor;
95          input name $ incrementt;
96          datalines;
```

NOTE: The data set WORK.INC_UNCOR has 8 observations and 2 variables.

NOTE: DATA statement used (Total process time):

```
real time      0.10 seconds
cpu time       0.00 seconds
```

```
103      ;
104      run;
105
106      /** increments for the corrected analysis. The order is same as the
107      above by replacing
108      surrogate with true exposure, that is, intercept, true exposure,
109      and the covariates **/
110
111      data inc_cor;
112          input name $ increment;
113          cards;
```

NOTE: SAS went to a new line when INPUT statement reached past the end of a line.

NOTE: The data set WORK.INC_COR has 8 observations and 2 variables.

NOTE: DATA statement used (Total process time):

```
real time      0.03 seconds
cpu time       0.02 seconds
```

```
120      ;
121      run;
122
123
124
125      /** call blinplus8 macro for each health outcome **/
126      %include "/udd/strui/sofie/blinplus8.sas";
127
128
129
130
131
132
133
134
135
136
137
138      title "For Current Wheeze";
139      %blinplus8( type      = GENMOD,
140                  valid     = valid,
141                  main_est  = b_wheeze,
142                  main_var  = cov_wheeze,
143                  increments1 = inc_uncor,
144                  increments2 = inc_cor,
145                  err_var   = surrno2 smokers gascook unvented sex age
146                  pet,
147                  true_var  = trueno2 smokers gascook unvented sex age
148                  pet,
149                  depend    = wheeze
150                  );
```

NOTE: Worksize = 1022976

NOTE: Symbol size = 262128

NOTE: IML Ready
NOTE: Closing WORK.COV_WHEEZE
NOTE: Exiting IML.
NOTE: The PROCEDURE IML printed page 10.
NOTE: PROCEDURE IML used (Total process time):
 real time 0.75 seconds
 cpu time 0.23 seconds

```
649
650
651     title "For Current Conjunctivitis";
652     %blinplus8(  type      = GENMOD,
653                 valid     = valid,
654                 main_est  = b_conj,
655                 main_var  = cov_conj,
656                 increments1 = inc_uncor,
657                 increments2 = inc_cor,
658                 err_var   = surrno2 smokers gascook unvented sex age
659                 true_var  = trueno2 smokers gascook unvented sex age
660                 depend   = conj
661                 );
NOTE: Worksize      = 1022976
NOTE: Symbol size = 262128
NOTE: IML Ready
NOTE: Closing WORK.COV_CONJ
```

The SAS System

NOTE: Exiting IML.
NOTE: The PROCEDURE IML printed page 11.
NOTE: PROCEDURE IML used (Total process time):
 real time 0.16 seconds
 cpu time 0.13 seconds

```
662
663     title "For Current Phlegm";
664     %blinplus8(   type      = GENMOD,
665                  valid     = valid,
666                  main_est  = b_phlegm,
667                  main_var  = cov_phlegm,
668                  increments1 = inc_uncor,
669                  increments2 = inc_cor,
670                  err_var   = surrno2 smokers gascook unvented sex age
671                  true_var  = trueno2 smokers gascook unvented sex age
672                  depend   = phlegm
673                  );
```

NOTE: Worksize = 1022976
NOTE: Symbol size = 262128
NOTE: IML Ready
NOTE: Closing WORK.COV_PHLEGM
NOTE: Exiting IML.
NOTE: The PROCEDURE IML printed page 12.
NOTE: PROCEDURE IML used (Total process time):
 real time 0.16 seconds
 cpu time 0.13 seconds

```
674
675
676
677     title "For total IgE";
678     %blinplus8(   type      = GENMOD,
679                  valid     = valid,
680                  main_est  = b_ige,
681                  main_var  = cov_ige,
682                  increments1 = inc_uncor,
683                  increments2 = inc_cor,
684                  err_var   = surrno2 smokers gascook unvented sex age
685                  true_var  = trueno2 smokers gascook unvented sex age
686                  depend   = ige
687                  );
```

NOTE: Worksize = 1022976
NOTE: Symbol size = 262128
NOTE: IML Ready
NOTE: Closing WORK.COV_IGE
NOTE: Exiting IML.
NOTE: The PROCEDURE IML printed page 13.
NOTE: PROCEDURE IML used (Total process time):
 real time 0.16 seconds
 cpu time 0.12 seconds

688

NOTE: SAS Institute Inc., SAS Campus Drive, Cary, NC USA 27513-2414

NOTE: The SAS System used:

real time	19.38 second
-----------	--------------