Example of PROC LOGISTIC matched pairs analysis by creating differences between members of a pair

- A conditional analysis for matched pairs can be conducted through the creation of differences of covariate values between members of a pair

`initial` = difference in initial skin grade values for each pair
`age` = difference in age value for each pair
`i_sex` = difference in sex value for each pair
`isexage` = difference in sex*age indicator value for each pair
`isexinit` = difference in sex*initial indicator value for each pair
`iageinit` = difference in age*initial indicator value for each pair
`trtsex` = difference in trt*sex indicator value for each pair
`trtinit` = difference in trt*initial indicator value for each pair
`trtage` = difference in trt*age indicator value for each pair

•Outcome variable uses only discordant pair information:
`improve=1` if (yes, no) for (new, placebo)
`improve=0` if (no, yes) for (new, placebo)
`improve=.` if else

```
proc logistic data=trial descending;
          model improve = initial age i_sex
              isexage isexinit iageinit
              trtsex trtinit trtage /
              selection=forward include=3 details;
      run;
```

Response Profiles

Response Profile

Ordered Value	Improve	Total Frequency
1	1	34
2	0	20

Score Statistics

Residual Chi-Square Test

Chi-Square	DF	Pr > ChiSq
4.9936	6	0.5446

Analysis of Effects Not in the Model

Effect	DF	Score Chi-Square	Pr > ChiSq
isexage	1	0.6593	0.4168
isexinit	1	0.0074	0.9312
iageinit	1	2.9194	0.0875
trtsex	1	0.2681	0.6046
trtinit	1	0.0121	0.9125
trtage	1	0.4336	0.5102

- Since there are 20 observations with the less prevalent response, this model can support 20/5 = 4 terms. Therefore, there are too many terms to rely on the residual score statistic to assess goodness of fit.
- However, the residual test ($p > 0.5$) and the individual tests (all $p > 0.08$) provide reasonable confidence that model fits adequately.

Maximum Likelihood Estimates

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	0.7024	0.3601	3.8053	0.0511
initial	1	1.0915	0.3351	10.6105	0.0011
age	1	0.0248	0.0224	1.2252	0.2683
i_sex	1	0.5312	0.5545	0.9176	0.3381

- The treatment effect is represented by intercept for this model.

Odds Ratios

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
Initial	2.979	1.545	5.745
Age	1.025	0.981	1.071
i_sex	1.701	0.574	5.043

• Note that LOGISTIC does not print the odds ratio for the intercept. Odds of improvement for those on treatment is $e^{0.7024} = 2.019$ times higher than for those on placebo.

• Consider the model where the intercept is the only term:

```
proc logistic data=trial descending;
   model improve = ;
run;
```

Treatment Effect Only Model

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	0.5306	0.2818	3.5457	0.0597

• Note that $e^{\beta} = e^{0.5306} = 1.70$