Optimal Data Analysis
Vol. 5 (May 29, 2016), 24-28

2155-0182/10/$3.00

ODA *vs*. Chi-Square: Describing Baseline Data from the National Pressure Ulcer Long-Term Care Study (NPULS)

Paul R. Yarnold, Ph.D.
Optimal Data Analysis, LLC

Chi-square analysis is often used to analyze data in contingency tables created by crossing two categorical variables, with at least one having three or more categories. Researchers report the associated omnibus (overall) *p* value to indicate the statistical reliability (not the strength) of the association between the variables. A statistically significant omnibus *p* value indicates two or more categories differ, but the exact structure of the inter-category difference(s) isn't explicit. Pairwise comparisons are needed to reveal the precise effect, but in practice this may be substituted for a non-statistical "eyeball analysis-based" summary of the data. In contrast, ODA models provide exact *p* values and an index of effect strength that is normed against chance and can be used to directly compare the classification accuracy achieved by alternative models. Furthermore, ODA models explicitly identify the structure of the omnibus effect, and an efficient optimal pairwise comparison methodology is used to ensure the statistical integrity of the model. These methods are illustrated for a sample of N = 2,420 adults at risk of developing a pressure ulcer.[1]

Table 1 presents categorical variables that were compared between four independent samples in the original study.[1] Groups were aggregated *a priori* on the basis of whether individuals were at risk for developing a pressure ulcer, or if they developed a pressure ulcer before or during the study: Group 1 is people at risk of developing a pressure ulcer at the start of the study; Group 2 is people who developed a new pressure ulcer during the study; Group 3 is people with an existing pressure ulcer at the start of the study; and Group 4 is people with both new and existing pressure ulcers.[1] Originally analyzed by chi-square[1] and evaluated here by a sequentially-rejective Sidak Bonferroni-type multiple comparisons procedure[2,3] to inhibit overfitting: between-sample differences in poor meal intake were not statistically significant; body mass index (BMI) differences were statistically reliable at the generalized (per-comparison) criterion; and statistically significant omnibus effects emerged at the experimentwise criterion ($p < 0.05$) for the remaining 11 variables (red text indicates lack of experimentwise significance).

2155-0182/10/$3.00

Table 1: Comparing Residents in Each of Four Study Groups

Patient Attribute	Category (%)	Group 1 N=1293	Group 2 N=457	Group 3 N=534	Group 4 N=136	χ^2 $p <$	ODA $p <$	ODA ESS
Gender	Male	321 (25)	130 (28)	212 (40)	58 (43)	0.001	0.0001	13.9
	Female	972 (75)	327 (72)	322 (60)	78 (57)			
Oral Nutritional Supplement	Yes	636 (49)	266 (58)	284 (53)	72 (53)	0.009	0.007	6.0
	No	657 (51)	191 (42)	250 (47)	64 (47)			
Enteral Feeding	Yes	216 (17)	76 (17)	145 (27)	50 (37)	0.001	0.0001	15.5
	No	1077 (83)	301 (83)	309 (73)	86 (63)			
Modular Products	Yes	81 (06)	74 (16)	111 (21)	43 (32)	0.001	0.0001	31.2
	No	1212 (94)	383 (84)	423 (79)	93 (68)			
Vitamin/Mineral Supplements	Yes	768 (59)	353 (77)	413 (77)	117 (86)	0.001	0.0001	23.4
	No	525 (41)	104 (23)	121 (23)	19 (14)			
Weight Loss Over 12-Week Study	Yes	646 (50)	263 (58)	249 (47)	66 (49)	0.006	0.021	5.3
	No	647 (50)	194 (42)	285 (53)	70 (51)			
BMI < 22 kg/m^2 at Study Start	Yes	558 (43)	216 (47)	268 (50)	62 (46)	0.043	0.018	5.3
	No	735 (57)	241 (53)	266 (50)	74 (54)			
Poor Meal Intake in First 4 Weeks	Yes	500 (39)	181 (40)	203 (38)	64 (47)	0.26	0.58	2.2
	No	793 (61)	276 (60)	331 (62)	72 (53)			
Cognitive Impairment	Yes	994 (77)	362 (79)	318 (60)	102 (75)	0.001	0.0001	15.6
	No	299 (23)	95 (21)	216 (40)	34 (25)			
Mobility Problems	Yes	1085 (84)	399 (87)	448 (84)	130 (96)	0.001	0.012	7.8
	No	208 (16)	58 (13)	86 (16)	6 (04)			
Incontinence	Yes	1197 (93)	420 (92)	459 (86)	129 (95)	0.001	0.0001	14.4
	No	96 (07)	37 (08)	75 (14)	7 (05)			
Hospitalized or Emergency Room	Yes	187 (14)	92 (20)	96 (18)	37 (27)	0.001	0.0007	9.7
	No	1106 (86)	365 (80)	438 (82)	99 (73)			
Died	Yes	56 (04)	42 (09)	58 (11)	19 (14)	0.001	0.0001	23.1
	No	1237 (96)	415 (91)	476 (89)	117 (86)			

Gender: The original article summarized the between-sample differences identified using chi-square for gender: "Group 4 had the highest percentage of males" (p. 1820).

The omnibus ODA model was: if Group is 1 or 2 then predict female; otherwise predict male: the effect was statistically significant at the experimentwise criterion, but was relatively weak (ESS = 13.9). Pairwise comparisons revealed no difference between Groups 1 and 2 (smallest proportion of males, $p < 0.14$, ESS = 3.6), or between Groups 3 and 4 (largest proportion of males, $p < 0.56$, ESS = 2.0).

2155-0182/10/$3.00

Oral nutritional supplement: The original article summarized between-sample differences identified using chi-square: "The extent of use of nutritional interventions was significantly different across the four study groups, with Group 1 residents having the lowest use" (p. 1822).

The omnibus ODA model was: if Group is 1 then predict no nutritional intervention; otherwise predict intervention: the effect was statistically significant at the generalized criterion but was only marginally stronger than chance (ESS = 6.0). A second ODA analysis revealed no difference between Groups 2, 3, and 4 ($p < 0.16$, ESS = 4.9).

Enteral feeding: The original article didn't discuss between-sample differences identified by chi-square for enteral feeding.

The omnibus ODA model was: if Group is 1 or 2 then predict no use of enteral feeding; otherwise predict the use of enteral feeding: the effect was statistically significant ($p < 0.0001$) but relatively weak (ESS = 15.5). Pairwise comparisons revealed no difference between Groups 1 and 2 ($p < 0.99$, ESS = 0.1), but Group 4 had higher use of enteral feeding than Group 3 when assessed at the generalized criterion ($p < 0.036$, ESS = 7.5).

Modular protein products: The original article didn't discuss between-sample differences identified by chi-square for use of modular protein products.

The omnibus ODA model was: if Group is 1 then predict no modular protein products; otherwise predict the use of modular protein products: the effect was statistically significant ($p < 0.0001$) and of moderate strength (ESS = 31.2). A second ODA analysis revealed a weak, statistically significant (generalized criterion) difference between Group 2, and Groups 3 and 4 ($p < 0.011$, ESS = 10.2). A final ODA analysis identified a weak, statistically significant (generalized criterion) difference between Groups 3 and 4 ($p < 0.008$, ESS = 9.9).

Vitamin/mineral supplements: The original article didn't discuss between-sample differences identified by chi-square for use of vitamin and mineral supplements.

The omnibus ODA model was: if Group is 1 then predict no vitamin and mineral supplements; otherwise predict the use of vitamin and mineral supplements: the effect was statistically significant ($p < 0.0001$) and nearly yielded moderate effect strength (ESS = 23.4). A second ODA analysis revealed a weak, statistically significant difference between Group 4, and Groups 2 and 3 ($p < 0.0030$, ESS = 11.0). A final ODA analysis found no difference between Groups 2 and 3 ($p < 0.99$, ESS = 0.1).

Weight loss over the 12-week study: The original article summarized the between-sample differences identified using chi-square: "There was a significant difference in the percentage of residents with weight loss during the study period among the four groups" (p. 1822).

The omnibus ODA model was: if Group is 2 then predict weight loss; otherwise predict no weight loss: the effect was statistically significant at the generalized criterion ($p < 0.021$) but was only marginally stronger than chance (ESS = 5.3). A second ODA analysis revealed no difference between Groups 1, 3, and 4 ($p < 0.29$, ESS = 2.6).

BMI < 22 kg/m^2 at study start: The original article summarized the between-sample differences identified using chi-square: "There was a significant difference in the percentage of residents with a BMI < 22 among Groups 1-4 with Group 3 having the largest percent" (p. 1822).

The omnibus ODA model was: if Group is 2 or 3 then predict BMI < 22; otherwise predict BMI $\geq$ 22: the effect was statistically significant at the generalized criterion, but was very weak (ESS = 5.3). Pairwise comparisons revealed no difference between Groups 1 and 4 ($p < 0.37$, ESS = 2.9), or between Groups 2 and 3 ($p < 0.65$, ESS = 0.9).

2155-0182/10/$3.00

Poor meal intake in first 4 weeks of the study: The original article summarized the lack of between-sample differences identified using chi-square: "There was no significant difference in meal intake among the groups" (p. 1822).

Using ODA no statistically significant effect was identified ($p < 0.58$, ESS = 2.2).

Cognitive impairment: The original article summarized the between-sample differences identified using chi-square: "A higher proportion of residents in Groups 1, 2, and 4 had some form of cognitive impairment compared with residents in Group 3" (p. 1822).

The omnibus ODA model was: if Group is 3 then predict no cognitive impairment; otherwise predict impairment: the effect was statistically significant at the experimentwise criterion but was relatively weak (ESS = 15.6). A second ODA analysis revealed no difference between Groups 1, 2, and 4 ($p < 0.41$, ESS = 2.6).

Mobility problems: The original article summarized the between-sample differences identified using chi-square: "There were 2,062 residents with mobility problems with a significant difference in at least two resident groups based on mobility; residents in Group 4 were the most immobile, followed by Group 2, Groups 1 and 3" (p. 1822).

The omnibus ODA model was: if Group is 1 or 3 then predict no mobility problems; otherwise predict mobility problems: the effect was statistically significant at the generalized criterion, and was very weak (ESS = 7.8). Pairwise comparisons revealed no difference between Groups 1 and 3 ($p < 0.99$, ESS = 0.1), however a smaller proportion of residents in Group 2 had mobility problems compared to residents in Group 4 ($p < 0.0065$, ESS = 15.2).

Incontinence: The original article summarized the between-sample differences identified using chi-square: "There were 2,206 residents who were incontinent of bowel or bladder with a significant difference in at least two resident groups based on incontinence. A lower proportion of residents in Group 3 had incontinence compared with residents in Groups 1, 2, and 4" (p. 1822).

The omnibus ODA model was: if Group is 3 then predict no incontinence; otherwise predict incontinence: the effect was statistically significant at the experimentwise criterion but was relatively weak (ESS = 14.4). A second ODA analysis revealed no difference between Groups 1, 2, and 4 ($p < 0.67$, ESS = 2.8).

Hospitalized or emergency room: The original article summarized the between-sample differences identified using chi-square: "A higher proportion of residents in Group 4 were hospitalized or had an emergency room visit than of residents in Groups 1, 2, or 3. There was no significant difference between residents in Groups 2 and 3, but Groups 2 and 3 had significantly higher hospitalizations and emergency room visits than Group 1" (p. 1822).

The omnibus ODA model was: if Group is 1 then predict no hospitalization or emergency room; otherwise predict hospitalization or emergency room: the effect was statistically significant at the experimentwise criterion but was relatively weak (ESS = 9.7). A second ODA analysis revealed no difference between Groups 2, 3, and 4 ($p < 0.19$, ESS = 5.9).

Mortality: The original article summarized the between-sample differences identified using chi-square: "A higher proportion of residents in Groups 3 and 4 died than of residents in Groups 1 and 2" (p. 1822).

The omnibus ODA model was: if Group is 1 then predict alive; otherwise predict dead: the effect was statistically significant at the experimentwise criterion and nearly yielded moderate effect strength (ESS = 23.1). A second ODA analysis revealed no difference between Groups 2, 3, and 4 ($p < 0.38$, ESS = 5.9).

Epilogue: A systematic and thorough examination of all effects identified in statistical analysis is needed to obtain solid understanding of between-sample/between-group differences. Of course, this is particularly important in applications that focus on between-group differ-

2155-0182/10/$3.00

ences—for example, in causal inference studies involving propensity score matching used to equate groups (samples) with respect to distributions of pre-intervention covariates, or dose-response research in which omnibus tests are followed by all-possible comparisons.[4,5]

References

[1]Horn SD, Bender SA, Bergstrom N, Cook AS, Ferguson ML, Rimmasch HL, Sharkey SS, Smout RJ, Taler GA, Voss AC (2002). *Journal of the American Geriatrics Society*, *50*, 1816-1825.

[2]Yarnold PR, Soltysik RC. *Optimal data analysis: Guidebook with software for Windows*. Washington, DC: APA Books, 2005.

[3]Yarnold PR, Soltysik RC (2016). *Maximizing predictive accuracy*. Chicago, IL: ODA Books.

[4]Linden A, Yarnold PR (2016). Using machine learning to assess covariate balance in matching studies. *Journal of Evaluation in Clinical Practice*. DOI: 10.1111/jep.12538

[5]Linden A, Yarnold PR, Nallamothu BK (2016). Using machine learning to model dose-response relationships. *Journal of Evaluation in Clinical Practice*.

Author Notes

The study analyzed de-individuated data and was exempt from Institutional Review Board review. No conflict of interest was reported.

Mail: Optimal Data Analysis, LLC
6348 N. Milwaukee Ave., #163
Chicago, IL 60646
USA