STATISTICAL ANALYIS PLAN WITH RESULTS

Screen to Save: A Colorectal Cancer Educational Intervention

NCT03907579

Date: July 8, 2020

ANALYSIS - BACKGROUND

For the Screen2Save intervention, subjects first take a survey, participate in the intervention and then take a final survey. The questions are the same on both surveys. All analysis was done using Stata/SE 15.1.

DID THE INTERVENTION CHANGE THE SURVEY RESULTS?

We are most interested in whether the subjects scored better (had more correct answers) in the post survey than the pre survey. Thus, for analysis purposes, each question was coded as a dichotomous Correct or Not Correct. If a subject did not submit a response for the question, it was coded as Not Correct (we may want to revisit this).

In order to test whether there was a difference in the Correct responses in the pre and post tests, McNemar's Test was used.

*The McNemar test is a non-parametric test for paired nominal data. It's used when you are interested in finding a change in proportion for the paired data. For example, you could use this test to analyze retrospective case-control studies, where each treatment is paired with a control. It could also be used to analyze an experiment where two treatments are given to matched pairs. This test is sometimes referred to as McNemar's Chi-Square test because the test statistic has a chi-square distribution. (*https://www.statisticshowto.com/mcnemar-test/)

When sample sizes are small, the chi-squared value may not be well-approximated by the chi-squared distribution. The exact p-value can be computed.

Thus, for each pair of pre and post questions, McNemar's Test was performed and both the Chi^2 and the exact P-value are shown. If the p-value is significant (<0.05), then the null hypothesis (H0 = the number of correct responses on the pre survey are the same as on the post survey) can be rejected.

HOW DID THE SURVEY RESPONSE CHANGE?

Although the primary interest is whether the intervention increased knowledge about Colorectal cancer and screening, understanding what the specific responses were to the surveys is also of interest. CrossTab tables help to both quantify and visualize the relationship between the Pre and Post responses. The Pre responses are shown in the rows while the Post responses are in the columns. The light green to dark blue color gradient highlights the low to high number of responses for each category.

Did the intervention increase the number of correct answers?

In order to look at the overall impact of the intervention, the number of correct responses was tabulated for each participant pre and post intervention. This data is shown in a CrossTab table. As with the individual questions, no response/null answers were coded as incorrect. In order to test of the number of correct responses was significantly different pre and post intervention, the Wilcoxon Signed-Rank Test was used. This is the non-parametric (i.e. not normally distributed) equivalent of a paired t-test.

SURVEY RESPONSES – WITH MCNEMAR'S CHI2 TEST STATISTIC AND EXACT P-VALUE AS WELL AS CROSSTAB

1. COLORECTAL CANCER STARTS IN WHAT PART OF THE BODY?

Correct Answer: Colon and/or rectum

McNemar's chi2(1) = 0.00 Prob > chi2 = 1.0000

Exact McNemar significance probability = 1.0000

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Starts in what body part				
Pre Starts in what body part	Null	Colon and or rectum	Small Intestine	Stomach	Grand Total
Null	2	4	1		7
Colon and or rectum	1	56	10	2	69
Small Intestine	1	8	10	1	20
Stomach		1	1	2	4
Grand Total	4	69	22	5	100

2.A STOOL TEST (FIT/FOBT) CHECKS YOUR STOOL (POOP) FOR?

Correct Answer: Blood

McNemar's chi2(1) = 3.20 Prob > chi2 = 0.0736

Exact McNemar significance probability = 0.1153

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Checks stool for					
Pre Checks stool for	Null	Blood	Fat	Polyps	Tumors	Grand Total
Null	3	6		4	1	14
Blood	2	60	2	1	1	66
Fat		2				2
Polyps	1	5		9	1	16
Tumors		1		1		2
Grand Total	6	74	2	15	3	100

3. WHICH OF THE FOLLOWING ARE RISK FACTORS FOR COLORECTAL CANCER? PLEASE CHECK ALL THAT ARE CORRECT.

Correct Answer: Count of 5

McNemar's chi2(1) = 21.78 Prob > chi2 = 0.0000

Exact McNemar significance probability = 0.0000

SIGNIFICANT, THERE IS A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Number of Risk Factors					
Pre Num..	1	2	3	4	5	Grand Total
0		1			3	4
1	1	1		3	8	13
2		5	4	3	7	19
3	2		2	1	6	11
4			1		8	9
5			1	3	40	44
Grand Total	3	7	8	10	72	100

4. IN GENERAL, A COLONOSCOPY SHOULD BE PERFORMED EVERY 10 YEARS STARTING AT AGE:

Correct Answer: 50

McNemar's chi2(1) = 13.00 Prob > chi2 = 0.0003

Exact McNemar significance probability = 0.0002

SIGNIFICANT, THERE IS A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Colonoscopy Starting age				
Pre Colonscopy ..	30	40	50	60	Grand Total
Null				1	1
30	3		1		4
40		4	11		15
50			79		79
60			1		1
Grand Total	3	4	92	1	100

5. IN GENERAL, A STOOL (POOP) TEXT (FIT OR FOBT) SHOULD BE DONE EVERY YEAR STARTING AT AGE:

Correct Answer: 50

McNemar's chi2(1) = 20.16 Prob > chi2 = 0.0000

Exact McNemar significance probability = 0.0000

SIGNIFICANT, THERE IS A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Stool test starting age					
Pre stool test start age	Null	30	40	50	60	Grand Total
Null	1			9		10
30		3	2	2		7
40	1	1	8	7		17
50		1	2	44		47
60			1	10	8	19
Grand Total	2	5	13	72	8	100

6. LYNCH SYNDROME IS A DISORDER THAT RUNS IN FAMILIES AND INCREASES MY CHANCES OF DEVELOPING COLORECTAL CANCER.

Correct Answer: True

McNemar's chi2(1) = 29.45 Prob > chi2 = 0.0000

Exact McNemar significance probability = 0.0000

SIGNIFICANT, THERE IS A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Lynch Syndrome increases risk			
Pre Lync..	Null	False	True	Grand Total
Null	5	2	23	30
False		1	17	18
True	3	1	48	52
Grand Total	8	4	88	100

7. THE EARLIER THAT COLORECTAL CANCER IS FOUND, THE GREATER MY CHANCES OF SURVIVAL.

Correct Answer: True

McNemar's chi2(1) = 2.00 Prob > chi2 = 0.1573

Exact McNemar significance probability = 0.5000

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Earlier found, greater survival		
Pre Earlier Found, greater survival	Null	True	Grand Total
Null	1	2	3
True		97	97
Grand Total	1	99	100

8. EVEN IF I HAVE NO SYMPTOMS, I MAY STILL HAVE COLORECTAL CANCER.

Correct Answer: True

McNemar's chi2(1) = 5.40 Prob > chi2 = 0.0201

Exact McNemar significance probability = 0.0352

SIGNIFICANT, THERE IS A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post No symptons			
Pre No s..	Null	False	True	Grand Total
Null	1		8	9
False			4	4
True	2	1	84	87
Grand Total	3	1	96	100

9. POLYPS ARE GROWTHS IN THE LINING OF THE COLON OR RECTUM THAT CAN DEVELOP INTO COLORECTAL CANCER.

Correct Answer: True

McNemar's chi2(1) = 0.00 Prob > chi2 = 1.0000

Exact McNemar significance probability = 1.0000

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Polyps are growths			
Pre Poly..	Null	False	True	Grand Total
Null	1		2	3
True	1	1	95	97
Grand Total	2	1	97	100

10. A COLONOSCOPY CAN BE USED TO FIND POLYPS IN THE COLON AND RECTUM.

Correct Answer: True

McNemar's chi2(1) = 1.00 Prob > chi2 = 0.3173

Exact McNemar significance probability = 1.0000

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Colonoscopy finds polyps		
Pre Colo..	Null	True	Grand Total
Null	1		1
False		1	1
True		98	98
Grand Total	1	99	100

11. A DIET HIGH IN RED MEATS AND PROCESSED MEATS (LUNCH MEATS, HOT DOGS) INCREASES MY CHANCES OF DEVELOPING COLORECTAL CANCER.

Correct Answer: True

McNemar's chi2(1) = 3.00 Prob > chi2 = 0.0833

Exact McNemar significance probability = 0.1460

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Meat increase cancer			
Pre Meat..	Null	False	True	Grand Total
Null	1		5	6
False		2	4	6
True	1	2	85	88
Grand Total	2	4	94	100

12. IT IS OK TO SKIP COLORECTAL CANCER SCREENING IF I DO NOT HAVE ANY SYMPTOMS.

Correct Answer: False

McNemar's chi2(1) = 0.33 Prob > chi2 = 0.5637

Exact McNemar significance probability = 0.7744

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Skip Screening			
Pre Skip ..	Null	False	True	Grand Total
Null	1	2		3
False		87	7	94
True		3		3
Grand Total	1	92	7	100

13. MY CHANCES OF DEVELOPING COLORECTAL CANCER ARE HIGHER IF SOMEONE IN MY IMMEDIATE FAMILY HAS IT OR HAS HAD IT.

Correct Answer: True

McNemar's chi2(1) = 4.50 Prob > chi2 = 0.0339

Exact McNemar significance probability = 0.0703

Chi2 is significant, but Exact p-value is not significant. Since the Incorrect responses (when added together) have a cellsize <4, then the Exact P-value should be used.

NOT SIGNIFICANT, THERE IS NOT A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Famial risk			
Pre Fami..	Null	False	True	Grand Total
Null	1		3	4
False			4	4
True		1	91	92
Grand Total	1	1	98	100

14. INCREASING MY PHYSICAL ACTIVITY WILL NOT LOWER MY CHANCES OF DEVELOPING COLORECTAL CANCER.

Correct Answer: False

McNemar's chi2(1) = 5.83 Prob > chi2 = 0.0158

Exact McNemar significance probability = 0.0241

SIGNIFICANT, THERE IS A SIGNIFICANT DIFFERENCE IN THE CORRECT RESPONSES PRE AND POST SURVEY.

	Post Physical activity NOT lower risk			
Pre Physi..	Null	False	True	Grand Total
Null	1	5	1	7
False		60	8	68
True		16	9	25
Grand Total	1	81	18	100

TOTAL NUMBER OF CORRECT ANSWERS

	Post Total Correct Answers							
Pre Total..	8	9	10	11	12	13	Yes	Grand T..
4		1						1
6			1	1				2
7	1		1			1		3
8	1	1		2	1	1		6
9			2		4			6
10	2	2	4	7	3	4		22
11			2		8	6		16
12				5	9	10		24
13				2	5	5		12
14			1		1	5		7
Yes							1	1
Grand To..	4	4	11	17	31	32	1	100

Wilcoxon Signed Rank Test

Sign	Obs	Sum ranks	Expected
Positive	25	1318	2420
Negative	55	3522	2420
Zero	20	210	210
All	100	5050	5050

unadjusted variance 84587.50
adjustment for ties -2388.38
adjustment for zeros -717.50
---
adjusted variance 81481.63
Ho: preTotalCorrect = postTotalCorrect
z = -3.861
Prob > |z| = 0.0001

SIGNIFICANT, THERE IS A SIGNIFICANT DIFFERENCE IN THETOTAL CORRECT RESPONSES PRE AND POST SURVEY.

SUMMARY STATISTICS

Pre Total Correct Answers

	Percentiles	Smallest		
1%	2.5	1		
5%	7	4		
10%	8	6	Obs	100
25%	10	6	Sum of Wgt.	100
50%	11		Mean	10.78
		Largest	Std. Dev.	2.195404
75%	12	14		
90%	13	14	Variance	4.819798
95%	14	14	Skewness	-1.335507
99%	14	14	Kurtosis	6.343168

Post Total Correct Answers

	Percentiles	Smallest		
1%	4.5	1		
5%	8.5	8		
10%	10	8	Obs	100
25%	11	8	Sum of Wgt.	100
50%	12		Mean	11.54
		Largest	Std. Dev.	1.719878
75%	13	13		
90%	13	13	Variance	2.95798
95%	13	13	Skewness	-2.702217
99%	13	13	Kurtosis	15.56301

Variable	Obs	Mean	Std. Dev.	Min	Max
preTotalCorrect	100	10.78	2.195404	1	14
postTotalCorrect	100	11.54	1.719878	1	13

The 1st column is Total Correct Answers Post – Pre. So, one person (Freq column) had 4 more correct answers on the pre test and one person had 6 more correct answers Post. In most cases (29), people had one more correct answer.

Post – Pre Total Correct	Freq.	Percent	Cum.
-4	1	1.00	1.00
-2	5	5.00	6.00
-1	19	19.00	25.00
0	20	20.00	45.00
1	29	29.00	74.00
2	9	9.00	83.00
3	11	11.00	94.00
4	2	2.00	96.00
5	3	3.00	99.00
6	1	1.00	100.00
Total	100	100.00	

INTENT

No analysis was done on the intent questions as they were only asked one time.