Introduction of Logistic regression and its use in health science research

Jason Liu

Brock University

Agenda

- What is logistic regression?
- Its application in health sciences research
- SAS codes

What is logistic regression?

Background

- Relationship
  - Describing the change of Y with X

- Regression model
  - $y_j = \beta_0 + \Sigma\beta_i x_j + e$ $\quad i = 1,2,\ldots,k$
  - *When i=1, simple linear regression*

$$\hat{y}_j = \beta_0 + \beta x$$

Both x and y $(-\infty, \infty)$

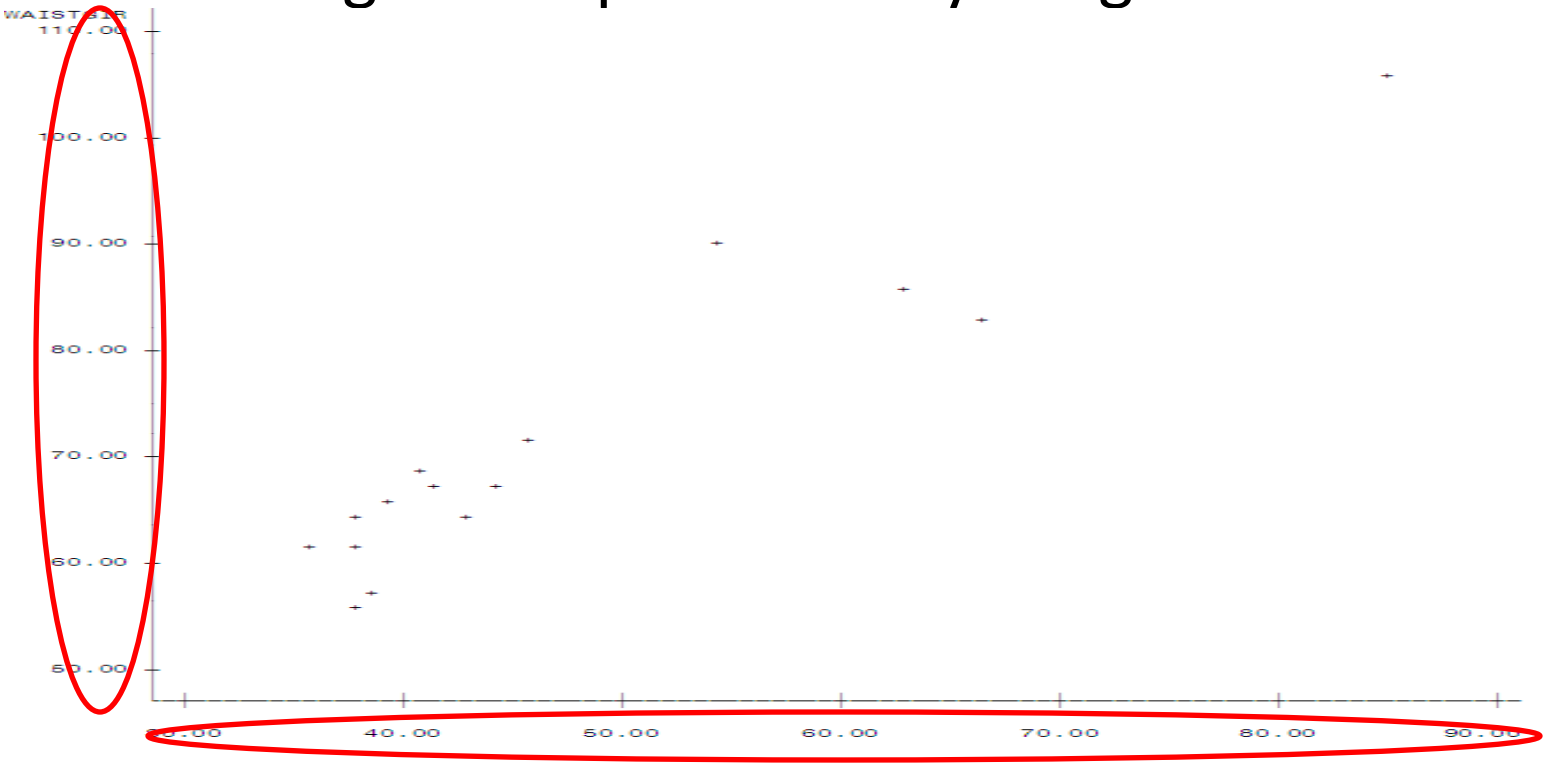
An example

Fifteen children with measurements of weight (kg), waist girth (cm), and overweight status

Obs	WEIGHT(kg)	WAIST(cm)	OWT
1	37.90	63.60	0
2	41.60	67.50	0
3	39.00	65.50	0
4	35.50	61.00	0
5	63.10	85.10	1
6	38.10	61.90	0
7	66.30	83.50	1
8	45.60	71.00	0
9	84.90	105.50	1
10	40.60	69.20	0
11	42.60	64.30	0
12	54.20	90.70	1
13	37.90	56.40	0
14	44.40	67.50	1
15	38.70	57.20	0

How is waist girth related to weight?

- Relationship
  - Positive vs. negative ?
  - Can waist girth be predicted by weight?



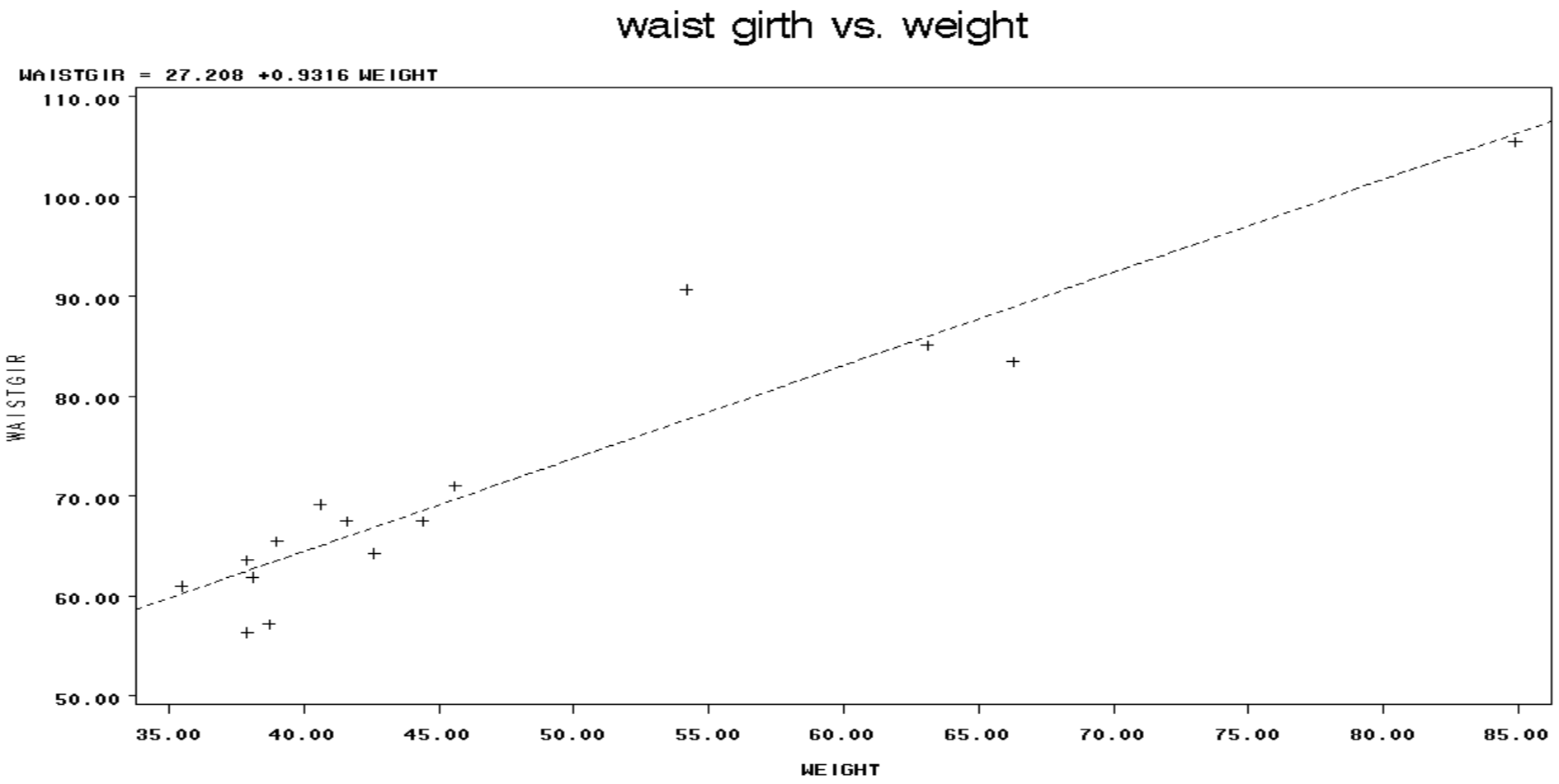
Waist girth

weight

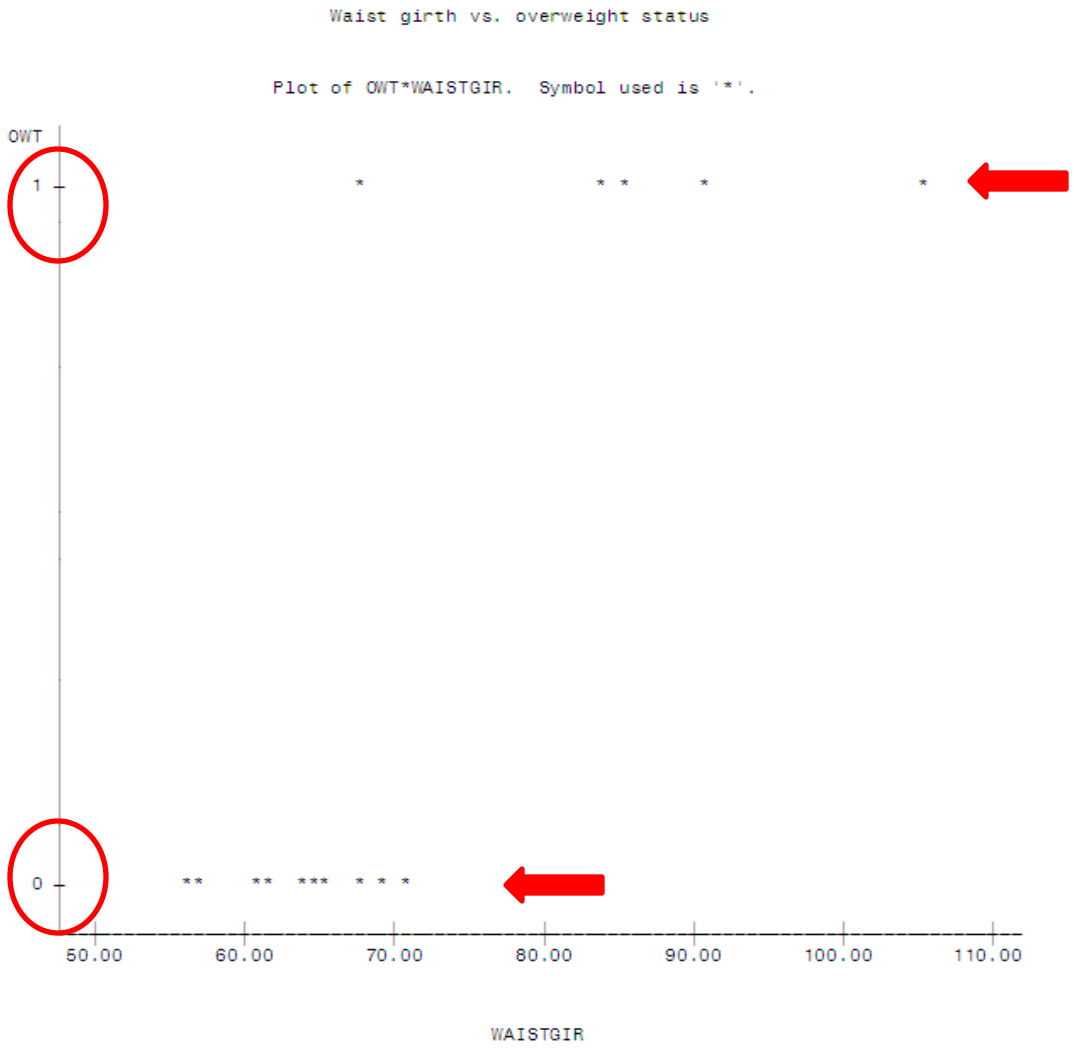
Can waist girth be predicted by weight?

- Regression equation:

Waist girth = 27.208 + 0.9316 weight



Waist girth vs. overweight status



Obs	WEIGHT(kg)	WAIST(cm)	OWT
1	37.90	63.60	0
2	41.60	67.50	0
3	39.00	65.50	0
4	35.50	61.00	0
5	63.10	85.10	1
6	38.10	61.90	0
7	66.30	83.50	1
8	45.60	71.00	0
9	84.90	105.50	1
10	40.60	69.20	0
11	42.60	64.30	0
12	54.20	90.70	1
13	37.90	56.40	0
14	44.40	67.50	1
15	38.70	57.20	0

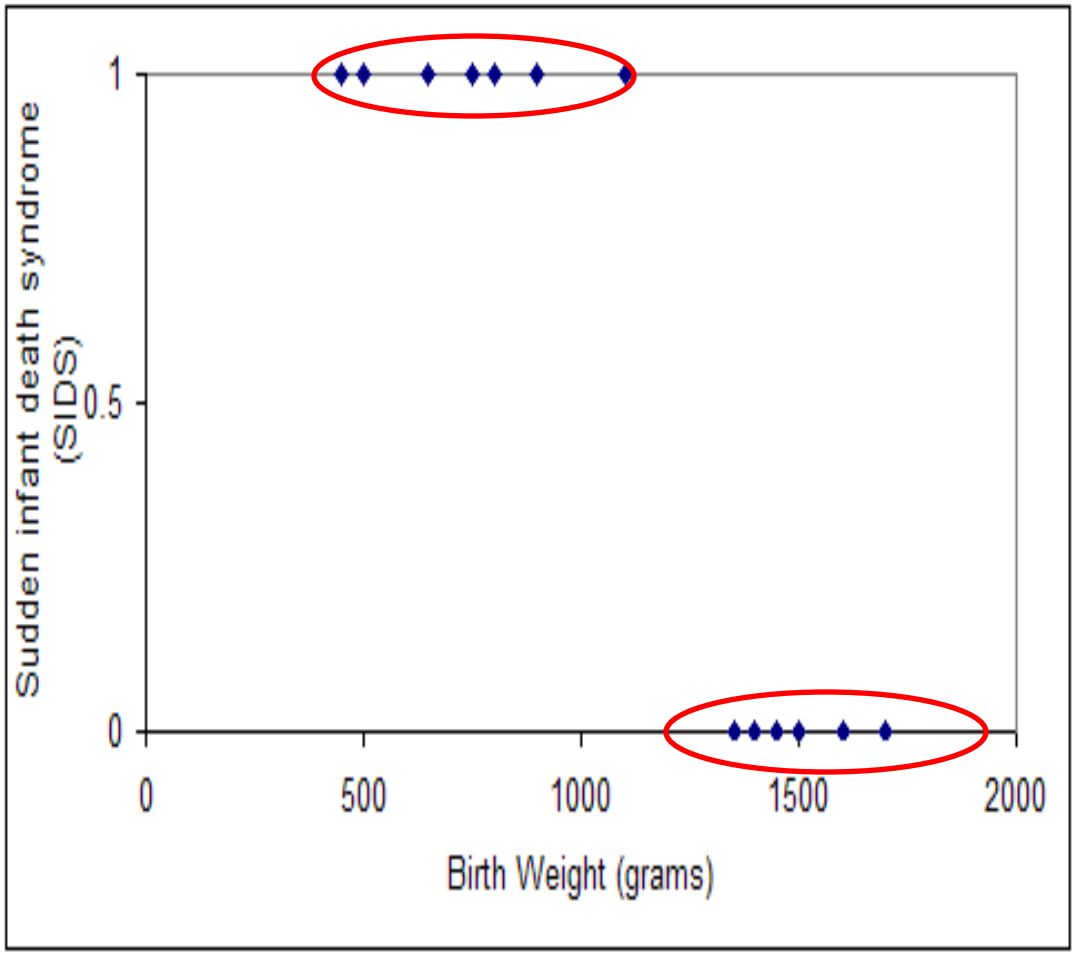
Dependent variable – dichotomous
Independent variable – continuous

- *Sudden infant death syndrome (SIDS) vs. birth weight (gms)*
  - *scatter diagram*

SIDS: *1* *yes*

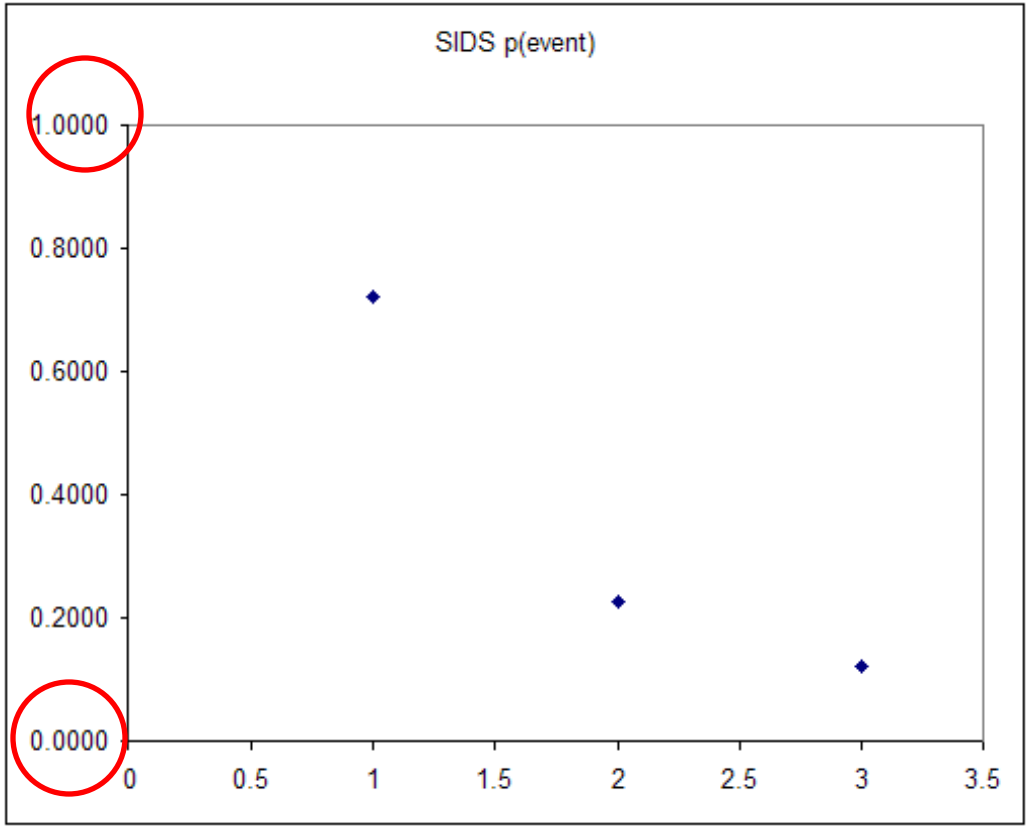
0 *no*

Birth wt: (-∞, ∞)



SIDS by birth wt category

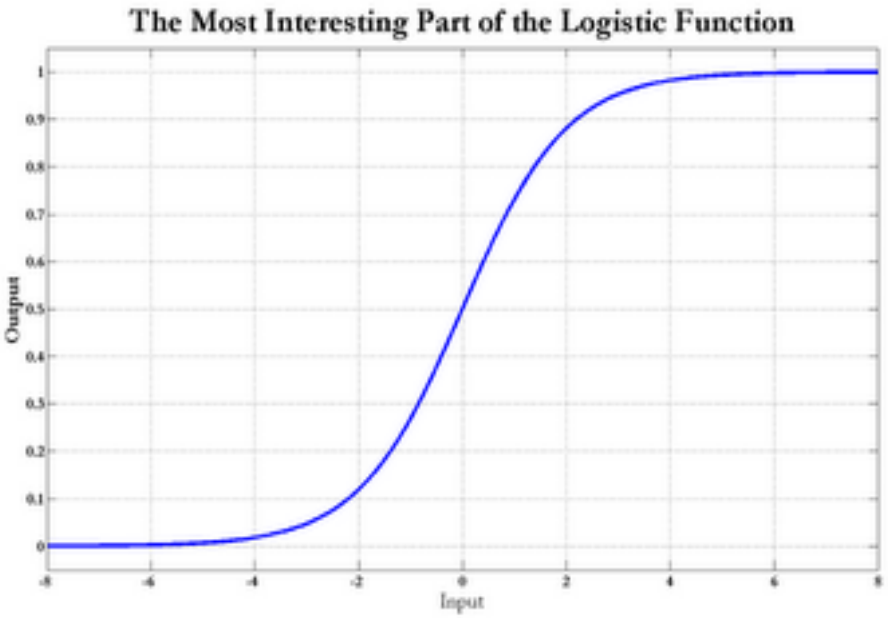
Birth Weigh (grams)	SIDS p(event)
0-950	0.7210
951-1350	0.2250
1351-175(	0.1200



Birth wt increase, the probability of SIDS decrease

What is logistic regression for?

- To examine the rela
  probability of an ev
  and other variable(



- Logistic function

$$p_i = \frac{1}{1+ e^{-(a+\beta x)}} \qquad i=1,2,\ldots,n. \qquad 0< p<1.$$

The odds in favour of the event

$$Pi/(1-pi) = e^{(a+\beta x)},$$

Taking the nature logarithm of each side of this equation,

$$\ln[pi/(1-pi)] = \ln[e^{(a+\beta x)}]$$
$$= a + \beta x$$

Regression model : $y = a + \beta x$

Its application in health sciences research

Sleep difficulty with childhood obesity

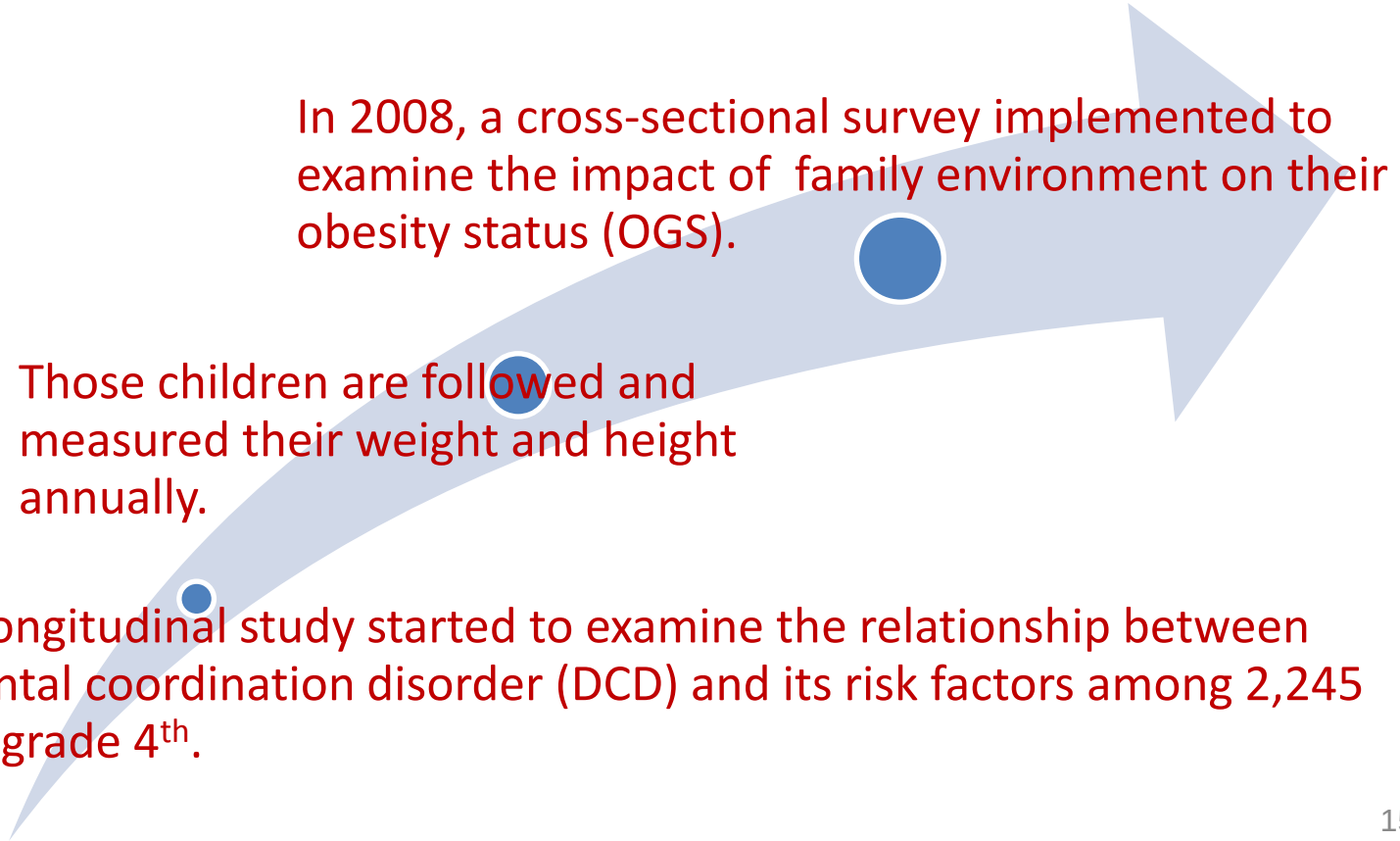
Aim: to determine if sleep difficulties are associated with childhood overweight and obesity.

- DV – overweight/obese (yes vs. no).
  - Probability of event (yes).

- IV - sleep difficulties.

Background of the study

Optimal Growth Study (OGS) is conducted among children in schools of Niagara Region.



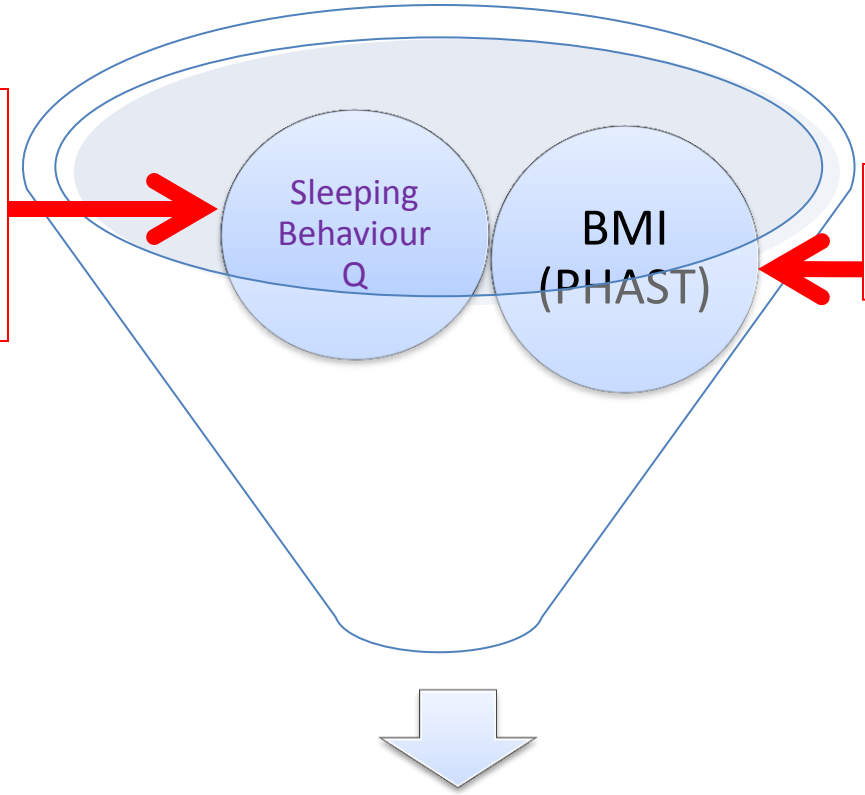
In 2008, a cross-sectional survey implemented to examine the impact of family environment on their obesity status (OGS).

Those children are followed and measured their weight and height annually.

In 2004, a longitudinal study started to examine the relationship between developmental coordination disorder (DCD) and its risk factors among 2,245 students in grade 4th.

Participants in this study

By parents and return to home class teacher (N=854)



Measured at school gym (N=2,205)

The final sample N = 606 (288 males, 318 females) after excluding those missing information

Sleeping Behavior questionnaire

WHAT TIME DOES YOUR CHILD USUALLY	Usual earliest time	Usual latest time
Go to bed on school days?		
Wake on school days?		
Go to bed on weekends?		
Wake on weekends?		

DOES YOUR CHILD:	Never	Sometimes	Often
Resist or object to going to bed?	1	2	3
Have trouble falling asleep?	1	2	3
Wake up at night?	1	2	3
Have nightmares?	1	2	3
Awaken screaming and inconsolable?	1	2	3
Snore or breathe loudly when asleep?	1	2	3
Move a lot or is he/she restless when asleep?	1	2	3
Frequently twitch or jerk his/her legs while asleep?	1	2	3
Seem sleepy or tired during the day?	1	2	3
Have trouble waking up in the morning?	1	2	3
Seem scared to sleep without parents nearby?	1	2	3

Covariates

- Age (yr)
- Gender (1 = male, 0 = female)

- Total physical activity score
- Total calories intake (kcal/day)
- Maternal education level (college or higher vs. less than college)

- Total sleeping hours (hrs).

Table 1. Characteristics of the participants and distribution of reported sleep behaviours by obesity status

Variable	Normal Mean (SD) n = 435	Overweight/obese Mean (SD) n = 171	P-value
Age (y)	11.3 (0.5)	11.3 (0.5)	0.6075
Male (%)	48.7	44.4	0.3411
Total physical activity score	19.1 (7.0)	17.2 (7.1)	0.0035
Total calories intake (Kcal/day)	2167.5 (1035.2)	2020.1 (1044.1)	0.1317
Body mass index (Kg/m^2)	17.8 (1.8)	24.5 (3.4)	<.0001
Waist circumference (cm)	65.7 (6.2)	84.0 (9.6)	<.0001
Mother's education Less than college^ (%)	42.7	49.1	0.0957
Sleep behavior problems (%)			
Wake up at night	44.5	55.6	0.0141
Snore when asleep	31.9	49.7	<.0001
Restless when asleep	52.8	63.2	0.0203
Jerk legs	26.2	29.2	0.4401
Resist going to bed	55.3	50.9	0.3281
Trouble falling asleep	66.1	60.8	0.2250
Have nightmares	28.4	33.3	0.2359
Awaken screaming and inconsolable	1.2	1.2	0.9811
Sleepy or tired during the day	59.6	66.1	0.1420
Trouble waking up in the morning	66.3	65.5	0.8538
Scared to sleep without parents nearby	15.8	18.7	0.3902
Total sleep behaviour problem score	16.2 (2.8)	16.8 (3.2)	0.0475
Sleep duration (hrs)	9.4 (0.7)	9.2 (0.7)	0.0002

^only 353 students had such information

SD – standard deviation

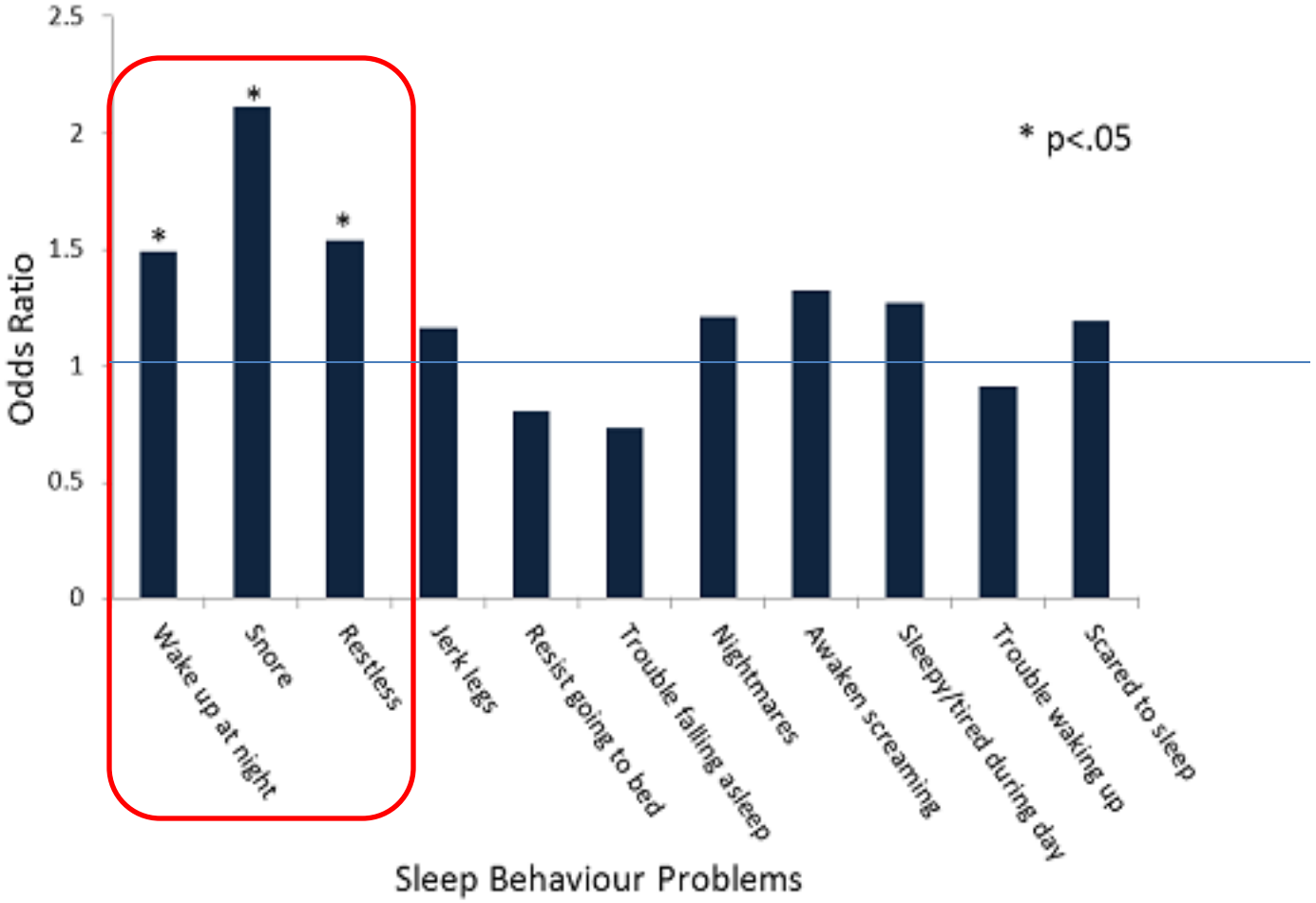


Figure 1: Odds ratio* of overweight/obese for sleep behaviour problems from logistic regression models

Note: *Adjusted for age, gender, and sleep duration

Table 2. Adjusted odds ratios* of overweight/obese by the combinations of sleep behaviour problems

	Sample Size (%)	OR	95% CI	P-value
None	117 (19.7)	1.00		
Wake-up	76 (12.8)	0.93	0.45, 1.94	0.8459
Snore	45 (7.6)	1.12	0.48, 2.65	0.7895
Restless	83 (14.0)	0.85	0.40, 1.80	0.6673
Wake-up & Snore	21 (3.5)	1.86	0.61, 5.66	0.4380
Wake-up & Restless	98 (16.5)	1.30	0.68, 2.55	0.2756
Snore & Restless	65 (10.9)	2.06	1.00, 4.22	0.0490
Wake-up, Snore, & Restless	90 (15.1)	3.43	1.80, 6.51	0.0002

*Adjusted for age, gender, total physical activity, total calories intake, and sleep duration

OR – odds ratio; CI – confidence interval.

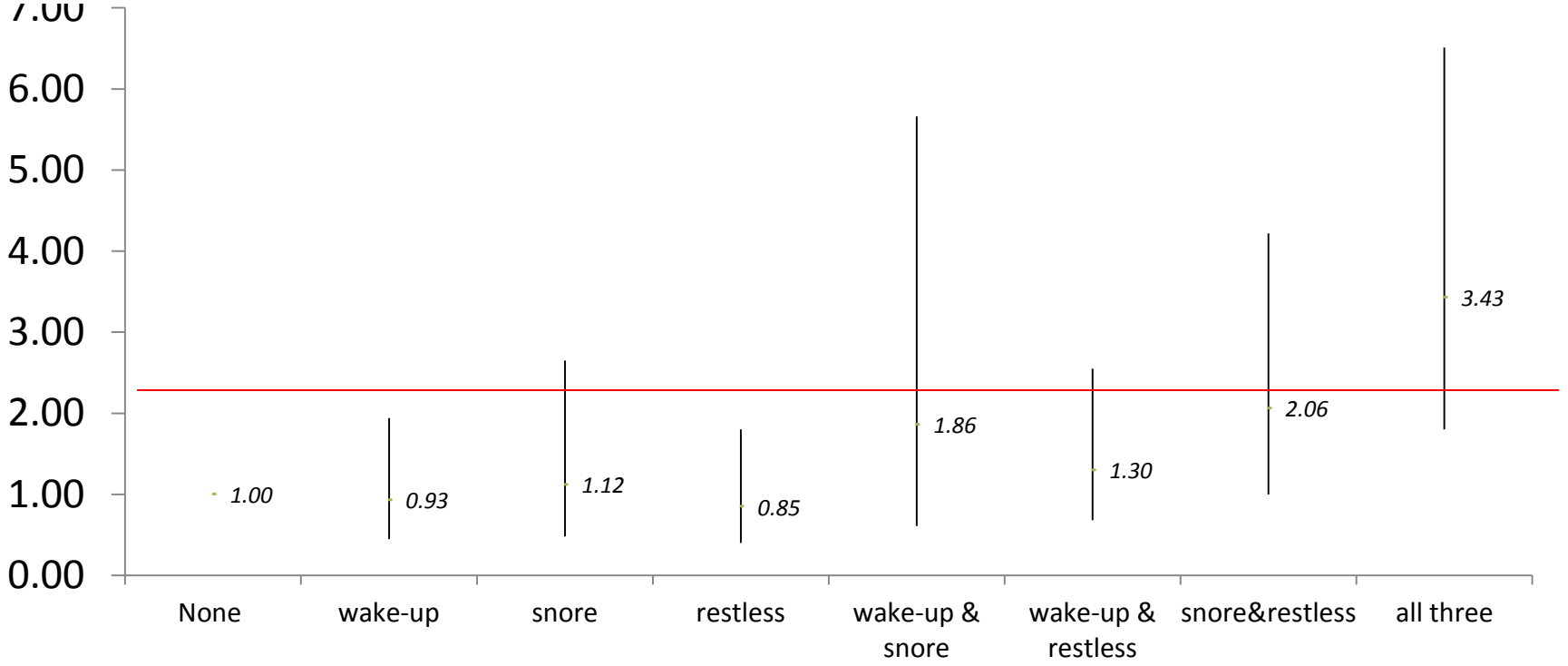
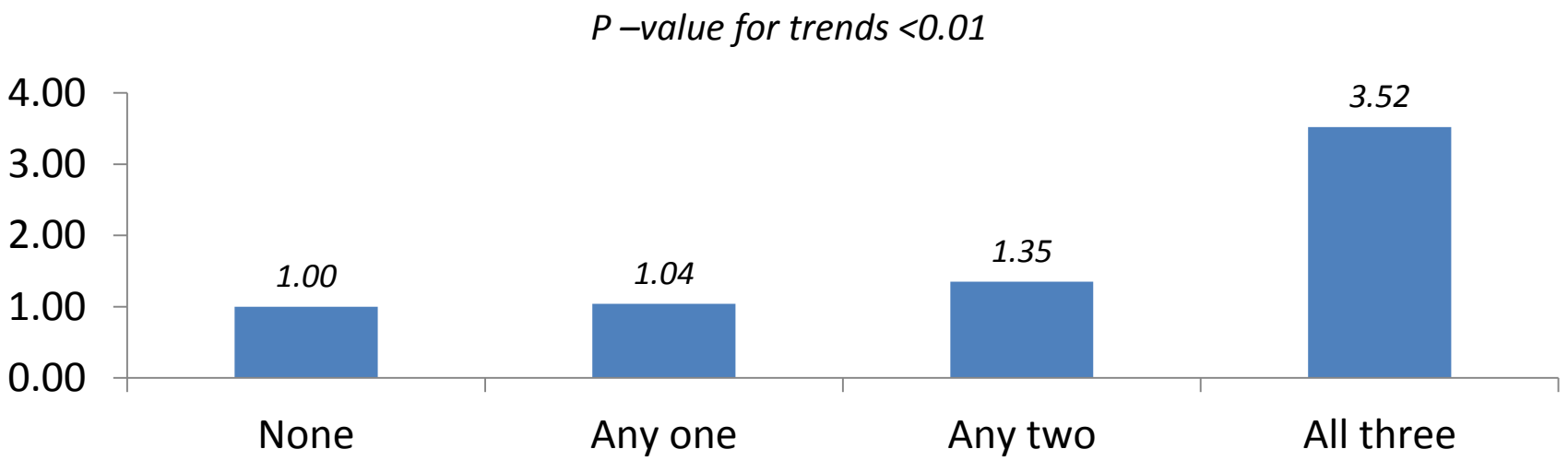


Table 3. Adjusted odds ratios* of overweight/obese by the number of sleep behaviour problems

	Any one sleep behaviour problem		Any two sleep behaviour problems		All three sleep behaviour problems		P-value for Trends
	OR	95% CI	OR	95% CI	OR	95% CI	
Model 1	0.97	0.51, 1.74	1.54	0.86, 2.76	3.53	1.87, 6.66	<.0001
Model 2	0.94	0.52, 1.69	1.58	0.88, 2.84	3.43	1.80, 6.52	<.0001
Model 3	1.04	0.46, 2.36	1.35	0.58, 2.10	3.52	1.42, 8.74	0.0043

Note: *Model 1, adjusted for age, gender, total physical activity, calorie intake; Model 2, further adjusted for sleep duration; Model 3, further adjusted for mother's education status.

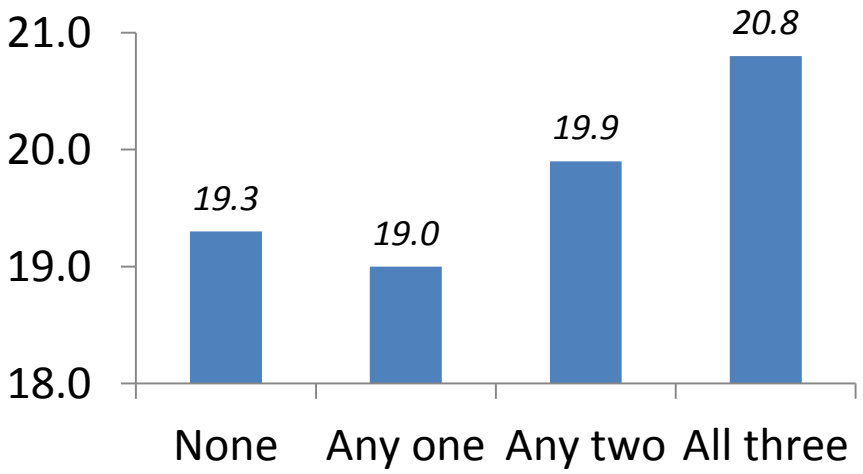
OR – odds ratio; CI – confidence interval.



Adjusted means of BMI and waist girth and adjusted prevalence of overweight/obese

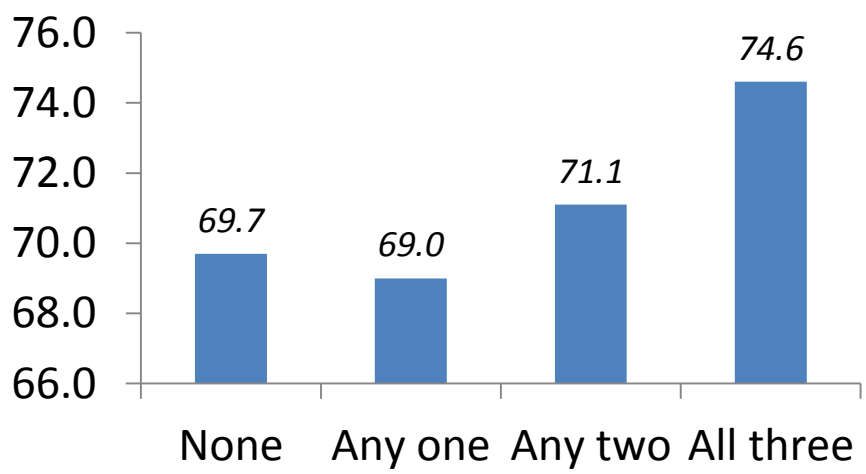
BMI

P –value for trends <0.01



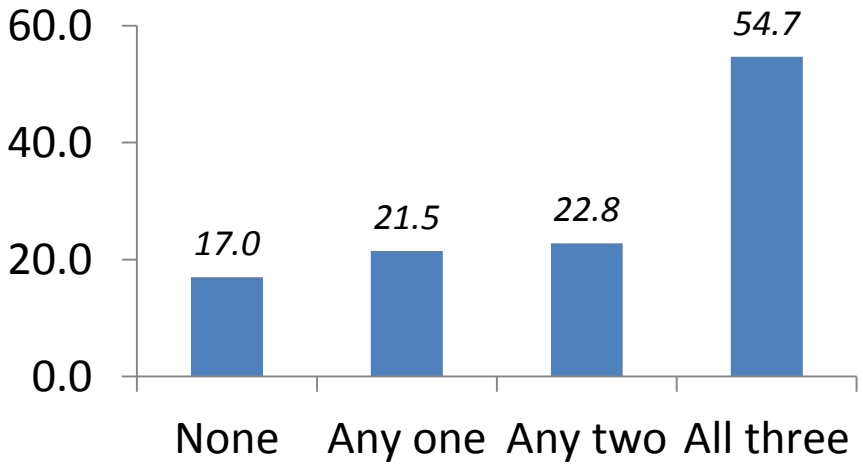
waist

P –value for trends <0.01



Owt/obese

P –value for trends <0.01



Adjusting for age, sex, total physical activity score, total calories intake, and sleep duration

In conclusion

Sleep difficulties are independently associated with overweight/obese status among children in Niagara Region.

SAS codes

Proc logistic procedure

Syntax

```
proc logistic descending;
    model DV =IVs
    covariates;
   title 'results for table 2';
run;
```

- Odds in favour of event (yes = 1).
- IVs and covariates can be dichotomous, ordinal, or continuous variables.

Proc catmod procedure

Syntax

```
proc catmod;
  weight wt;
  direct x1 x2;
  model r=x1 x2;
Run;
```

- Contingency table data

Population Profiles

Sample	Heat	Soak	Sample Size
1	7	1	10
2	7	1.7	17
3	7	2.2	7
4	7	2.8	12
5	7	4	9
6	14	1	31
7	14	1.7	43
8	14	2.2	33
9	14	2.8	31
10	14	4	19
11	27	1	56
12	27	1.7	44
13	27	2.2	21
14	27	2.8	22
15	27	4	16
16	51	1	13
17	51	1.7	1
18	51	2.2	1
19	51	4	1

Response Profiles

Response	Y
1	0
2	1

The CATMOD Procedure

Response Frequencies

	Response Number	
Sample	1	2
1	10	0
2	17	0
3	7	0
4	12	0
5	9	0
6	31	0
7	43	0
8	31	2
9	31	0
10	19	0
11	55	1
12	40	4
13	21	0
14	21	1
15	15	1
16	10	3
17	1	0
18	1	0
19	1	0

Proc genmod procedure

```
data drug;
   input drug$ x r n @@;
   datalines;
   A  .1   1  10   A  .23  2  12   A  .67  1   9
   B  .2   3  13   B  .3   4  15   B  .45  5  16   B  .78  5  13
   C  .04  0  10   C  .15  0  11   C  .56  1  12   C  .7   2  12
   D  .34  5  10   D  .6   5   9   D  .7   8  10
   E  .2  12  20   E  .34 15  20   E  .56 13  15   E  .8  17  20
   ;
run;
```

Syntax

```
 proc genmod;
    class drug;
    model r/n = x drug / dist = bin
                         link = logit
                              lrci;
run;
```

Thank you!