

# Clusters of Injuries and Chronic Health Conditions and their changes between baseline and follow-up for the kindergarten cohort

Jack Chen MBBS PhD MBA(Exec),  
Lixin Ou MBA MPH

Simpson Centre for Health Services Research  
The University of New South Wales

# The LSAC study

- ✦ One of the most well-designed, funded and managed longitudinal studies in the world;
- ✦ A wealth of data and framework;
- ✦ Great opportunities and great challenges in understanding the data



# The ideas drive this study

- ✦ The study asked a battery of questions (11) about the chronic health conditions of the participating children at both baseline and follow-up for K-cohort;
- ✦ The study also asked a group of questions regarding the injuries suffered in the past 12 months (10 items at baseline and follow-up);
- ✦ We also collected a great deal of predicting variables (ecological, social, family, personal, school, etc)

# Questions

- ✦ Are there any underlying distinctive groups behind the chronic health conditions and injuries, separately?
- ✦ If yes, what are the relationships between health condition groups and injuries groups at the baseline ?
- ✦ Do these groups change over the time ?
- ✦ More importantly, what are the predictors of these groups?



*Children ongoing health problems at baseline*

4/5 – f2f b36 – sc had wheezing or whistling

4/5 – f2f b37 – sc had illness with wheezing

4/5 – f2f b44 – sc had bronchiolitis

4/5 – f2f b46.1 – hearing problems

4/5 – f2f b46.2 – vision problems

4/5 – f2f b46.6 – frequent headaches

4/5 – f2f b46.7 – ear infections

4/5 – f2f b46.8 – other infections

4/5 – f2f b46.9 – food or digestive allergies

4/5 – f2f b46.10 – other illnesses

4/5 – f2f b46.11 – other physical disabilities

## *10 baseline measures of injuries in past 12 months*

4/5 –f2f b50.1 –injury –broken bones

4/5 –f2f b50.2 –injury –burn or scald

4/5 –f2f b50.3 –injury –dislocation

4/5 –f2f b50.4 –injury –sprain or strain

4/5 –f2f b50.5 –injury –cut or scrape

4/5 –f2f b50.6 –injury –concussion

4/5 –f2f b50.7 –injury –internal (not head)

4/5 –f2f b50.8 –injury –dental

4/5 –f2f b50.9 –injury –accidental poisoning

4/5 –f2f b50.10 –injury –other

## *11 follow-up measures of children ongoing health conditions*

**6/7 – f2f b8.1 – sc had illness with wheezing**

**6/7 – f2f b9.1 – sc diagnosed with asthma**

**6/7 – f2f b14.1 – sc's sleeping patterns a problem**

**6/7 – f2f b11.2.1 – eczema**

**6/7 – f2f b11.2.2 – attention deficit disorder**

**6/7 – f2f b11.2.3 – frequent headaches**

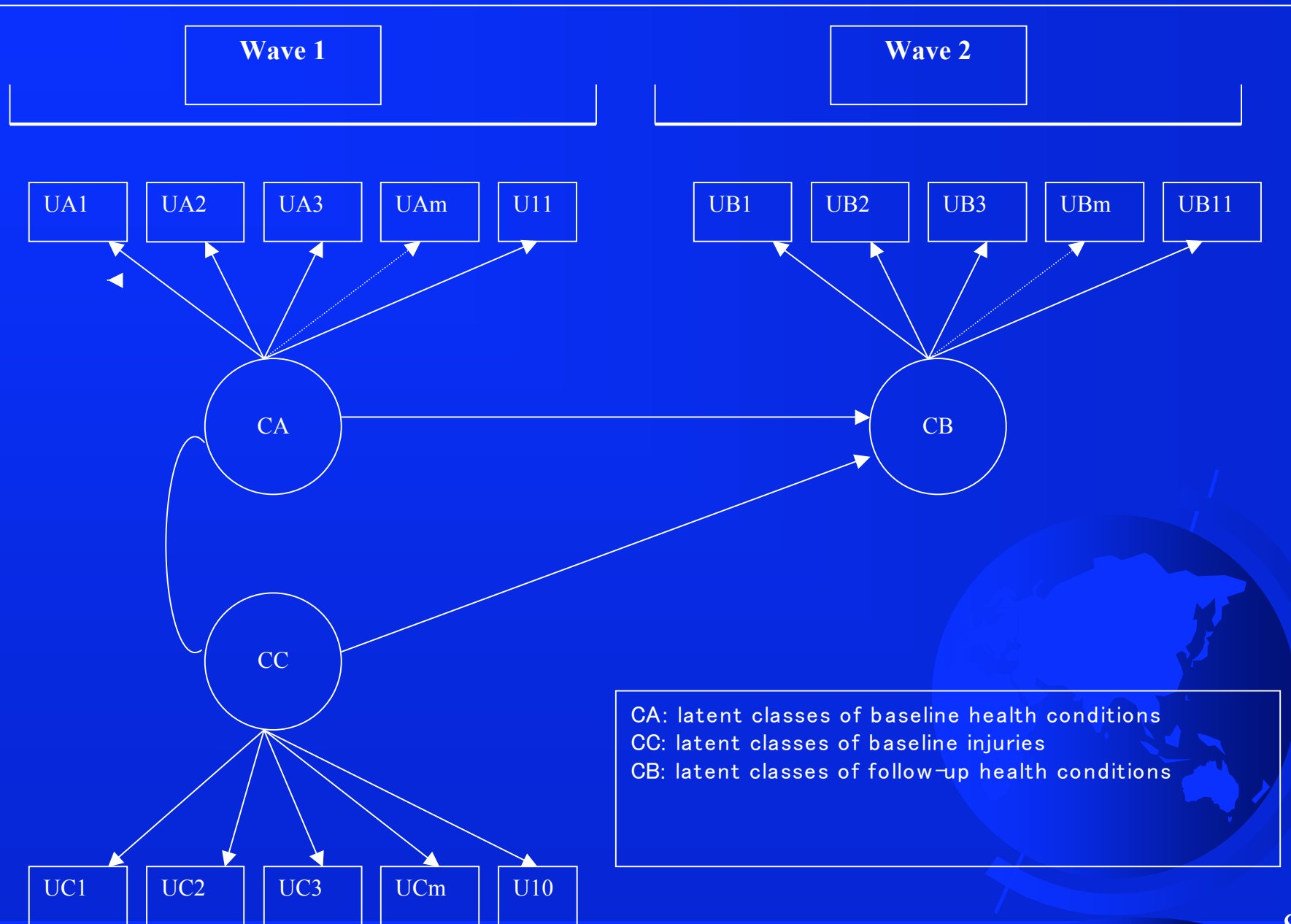
**6/7 – f2f b11.2.4 – diarrhoea/collitis**

**6/7 – f2f b11.2.5 – ear infections**

**6/7 – f2f b11.2.6 – other infections**

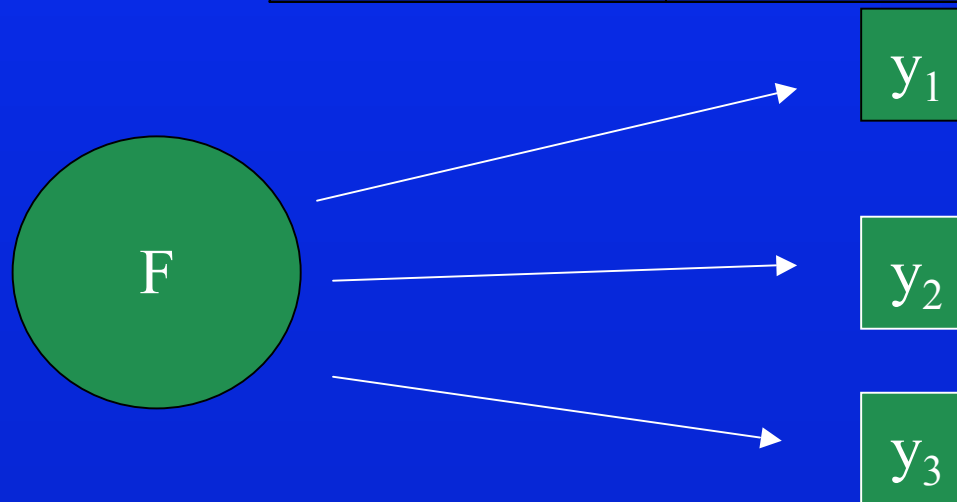
**6/7 – f2f b11.2.7 – food or digestive allergies**

**6/7 – f2f b11.2.9 – other illnesses**



# Introduction: types of Latent Variable Models: traditional taxonomy

		Observed Variables	
		Continuous	Categorical
Latent Variable	Continuous	Factor analysis	Latent trait Analysis
	Categorical	Latent profile analysis	Latent class analysis

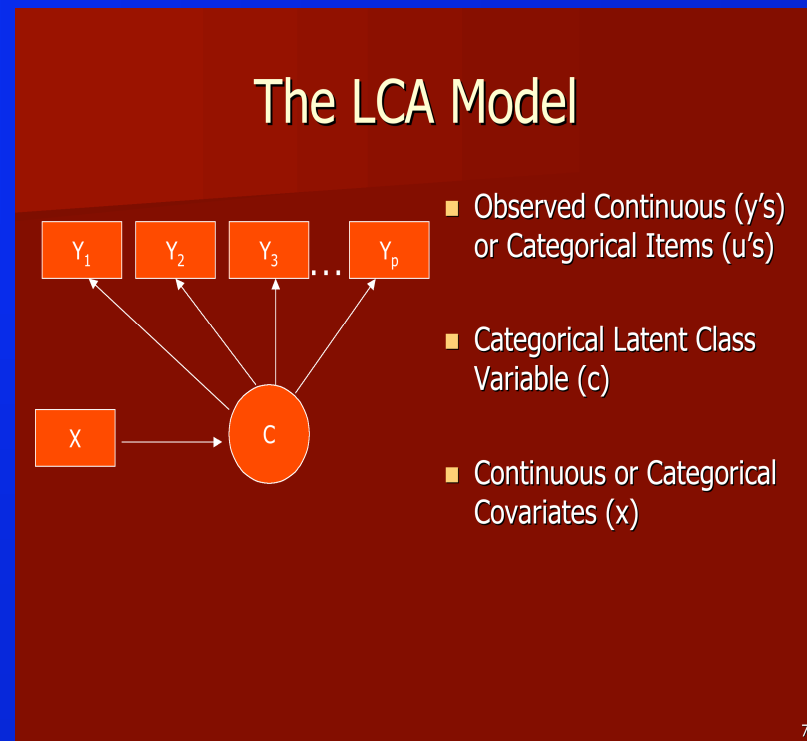


# Latent Variable Models: Modern Taxonomy

Model	Latent variable	Observed variable
<b>Confirmatory and Exploratory Factor Analysis</b>	<b>Continuous</b>	<b>Continuous or categorical</b>
<b>Latent Profile Analysis</b>	<b>Categorical</b>	<b>Continuous</b>
<b>Latent Trait Analysis (including IRT)</b>	<b>Continuous</b>	<b>Categorical</b>
<b>Latent Class Analysis (LCA)</b>	<b>Categorical</b>	<b>Categorical or Continuous</b>

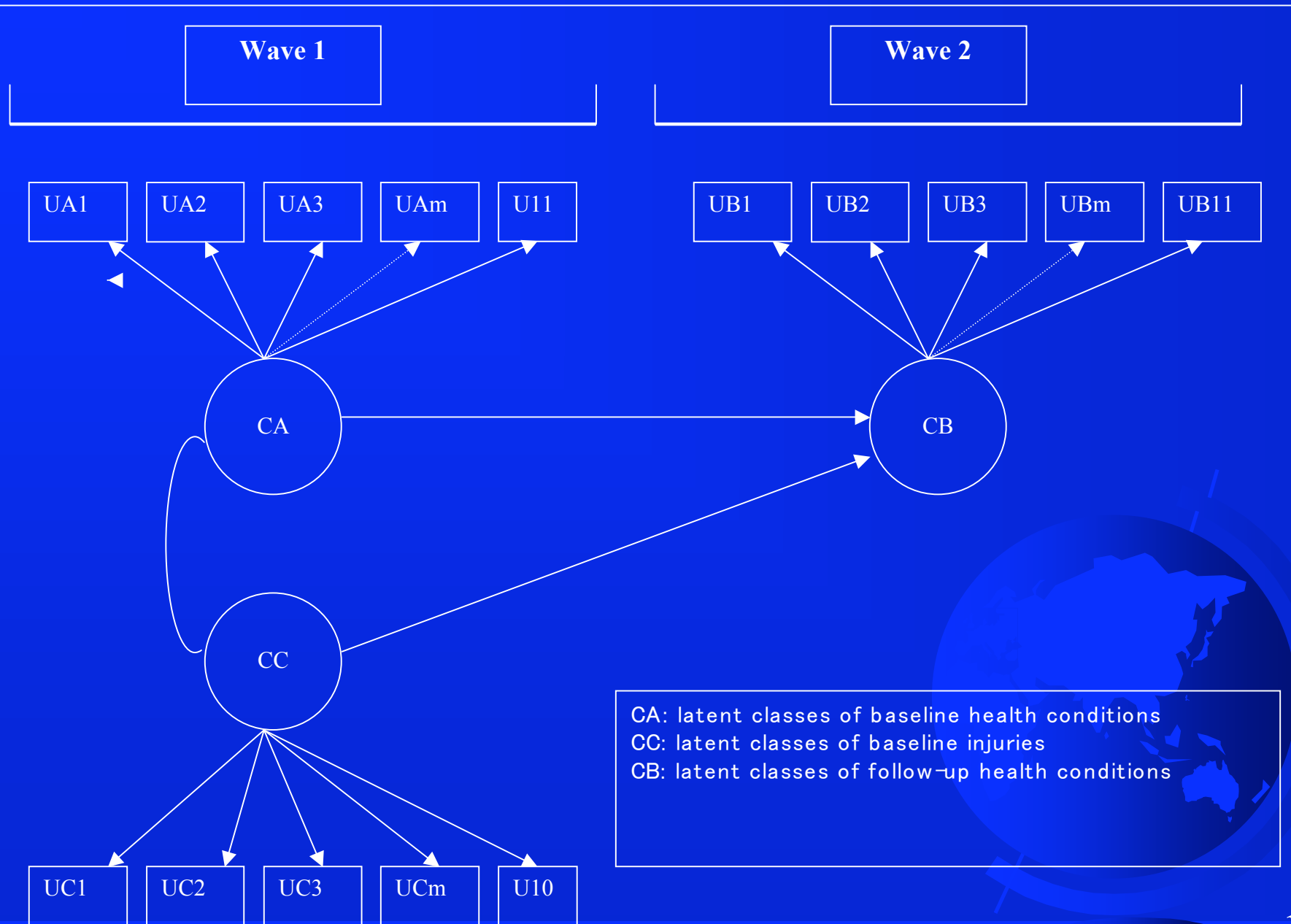
# What is a LCA model ? My definition

A model that tries to identify the potential latent classes (C) underlying the whole population based on the observed outcome measures (Ys) and population characteristics (Xs).



# LCA Purpose

- ✦ **Use patterns of responses to observed variables to:**
  - **identify the number of underlying classes;**
  - **classify each individual into one class;**
  - **determine class size; and**
  - **the distribution of observed variables within each class**



# How do we estimate such a model ?

- ✦ The second general structure equation framework (Muthen & Muthen)
- ✦ Software: Mplus ( version 5 in Nov. 2007)



# Strategy

- ✦ Estimate the number of the classes for chronic health conditions and injuries, separately;
- ✦ Fit a comprehensive model LCA;
- ✦ Predicted the Class membership and save it to the data;
- ✦ Multiple imputation using ICE program in Stata
- ✦ Estimate the predictive model in Stata using multiple imputed dataset



# What can MPlus v 5 do?

- ✦ To decide how many classes
- ✦ To estimate the membership of class for each and every child
- ✦ To allow the complex relationships among different latent classes
- ✦ To save the membership of the class for further analysis
- ✦ To deal with complex and missing values properly



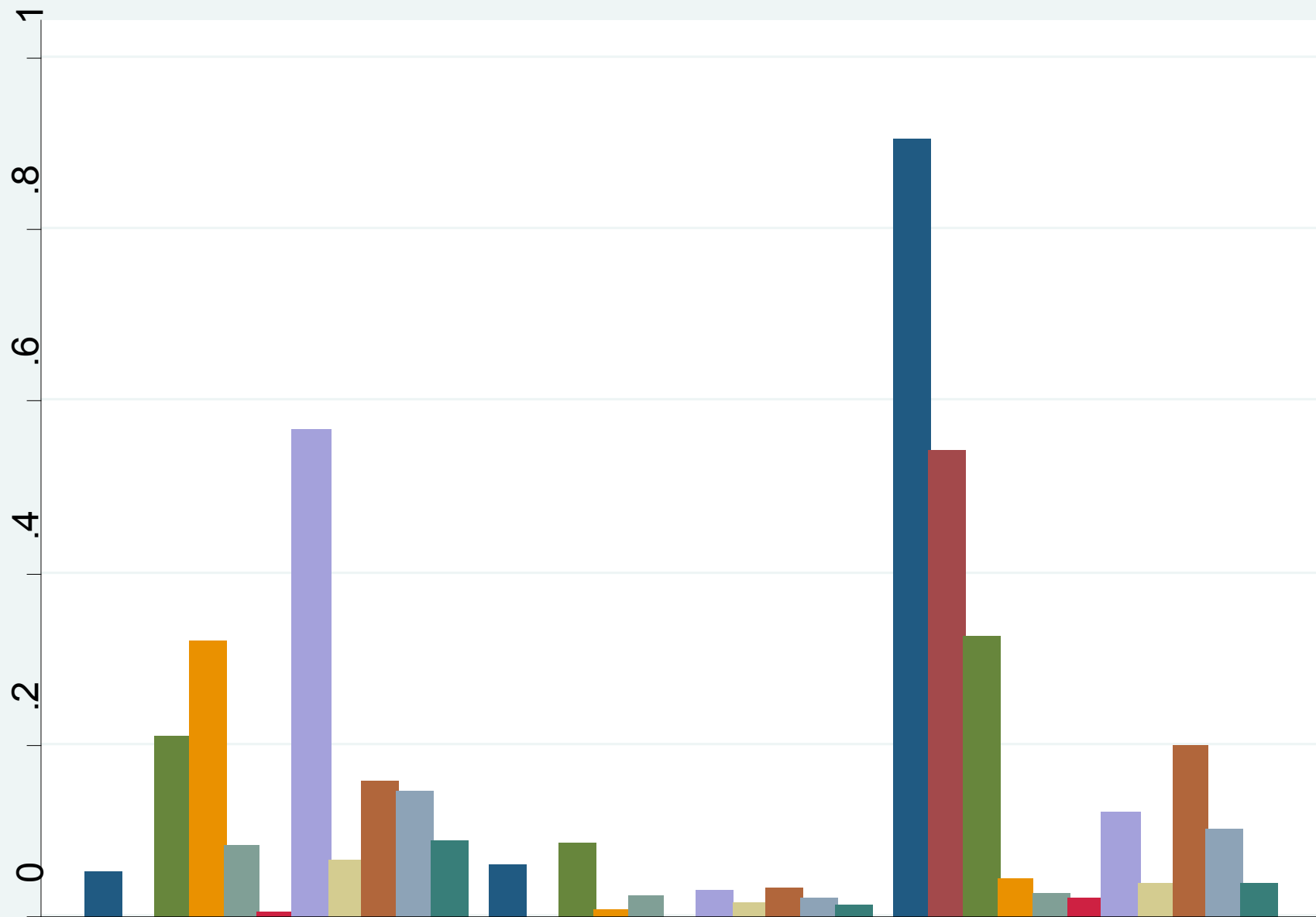
# How many latent classes did data suggest?

✦ Baseline : Health conditions 3; Injury: 2

✦ Follow-up: Health conditions 3; Injury : 1



	The Latent classes of chronic health conditions at W 1						Naïve sample estimate		Complex sample estimate
	Class 1		Class 2		Class 3		n	%	%
	n	%	n	%	n	%			
4/5 – f2f b36 – sc had wheezing or whistling	237	5.1%	3392	6.1%	1344	90.7%	4973	28.9%	29.7%
4/5 – f2f b37 – sc had illness with wheezing	237	0.0%	3391	0.0%	1335	54.4%	4963	14.6%	15.2%
4/5 – f2f b44 – sc had bronchiolitis	238	21.0%	3371	8.5%	1328	32.6%	4937	15.6%	16.0%
4/5 – f2f b46.1 – hearing problems	239	32.2%	3397	0.8%	1347	4.4%	4983	3.3%	3.3%
4/5 – f2f b46.2 – vision problems	239	8.4%	3397	2.6%	1347	2.8%	4983	2.9%	2.9%
4/5 – f2f b46.6 – frequent headaches	239	0.4%	3397	0.0%	1347	2.2%	4983	0.6%	0.6%
4/5 – f2f b46.7 – ear infections	239	56.9%	3397	2.8%	1347	12.0%	4983	7.9%	8.0%
4/5 – f2f b46.8 – other infections	239	6.7%	3397	1.5%	1347	3.8%	4983	2.3%	2.3%
4/5 – f2f b46.9 – food or digestive allergies	239	15.9%	3397	3.1%	1347	20.0%	4983	8.3%	8.4%
4/5 – f2f b46.10 – other illnesses	239	14.6%	3397	2.1%	1347	10.3%	4983	5.0%	5.0%
4/5 – f2f b46.11 – other physical disabilities	239	8.8%	3397	1.3%	1347	3.7%	4983	2.3%	2.4%



Class 1

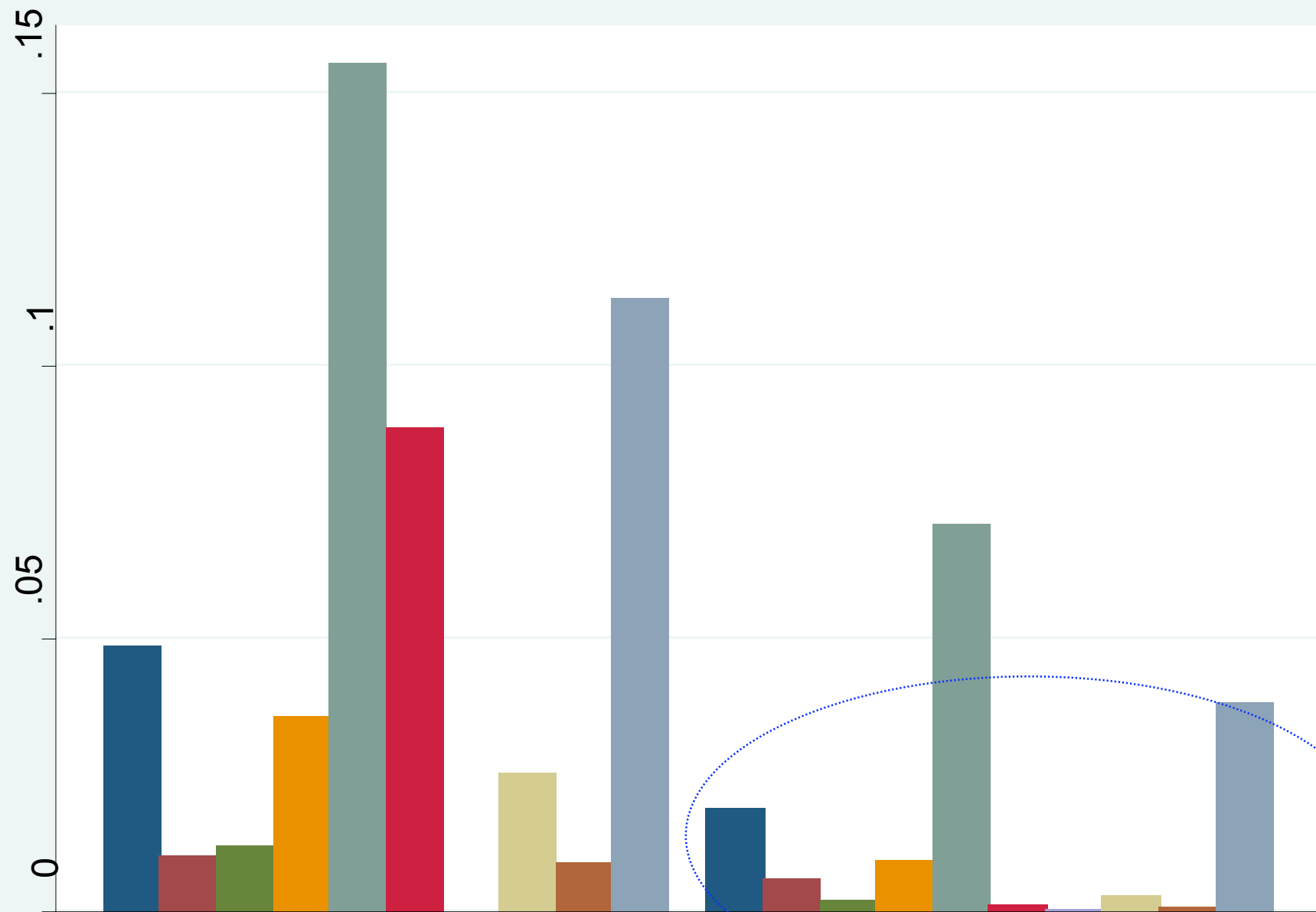
Class 2

Class 3

Latent classes of chronic health conditons at w1

	The Latent classes of chronic health conditions at W 1						Complex sample estimate	
	Class 1		Class 2		Native sample estimate			
	n	%	n	%	n	%		
4/5 –f2f b50.1 –injury – broken bones	676	4.9%	4272	1.9%	4948	2.3%	2.3%	
4/5 –f2f b50.2 –injury –burn or scald	676	1.0%	4272	0.6%	4948	0.7%	0.7%	
4/5 –f2f b50.3 –injury – dislocation	676	1.2%	4272	0.2%	4948	0.3%	0.3%	
4/5 –f2f b50.4 –injury – sprain or strain	676	3.6%	4272	1.0%	4948	1.3%	1.3%	
4/5 –f2f b50.5 –injury –cut or scrape	676	15.5%	4272	7.1%	4948	8.3%	8.4%	
4/5 –f2f b50.6 –injury – concussion	676	8.9%	4272	0.1%	4948	1.3%	1.4%	
4/5 –f2f b50.7 –injury – internal (not head)	676	0.0%	4272	0.0%	4948	0.0%	0.0%	
4/5 –f2f b50.8 –injury – dental	676	2.5%	4272	0.3%	4948	0.6%	0.8%	
4/5 –f2f b50.9 –injury – accidental poisoning	676	0.9%	4272	0.1%	4948	0.2%	0.2%	
4/5 –f2f b50.10 –injury – other	676	11.2%	4272	3.8%	4948	4.8%	5.2%	



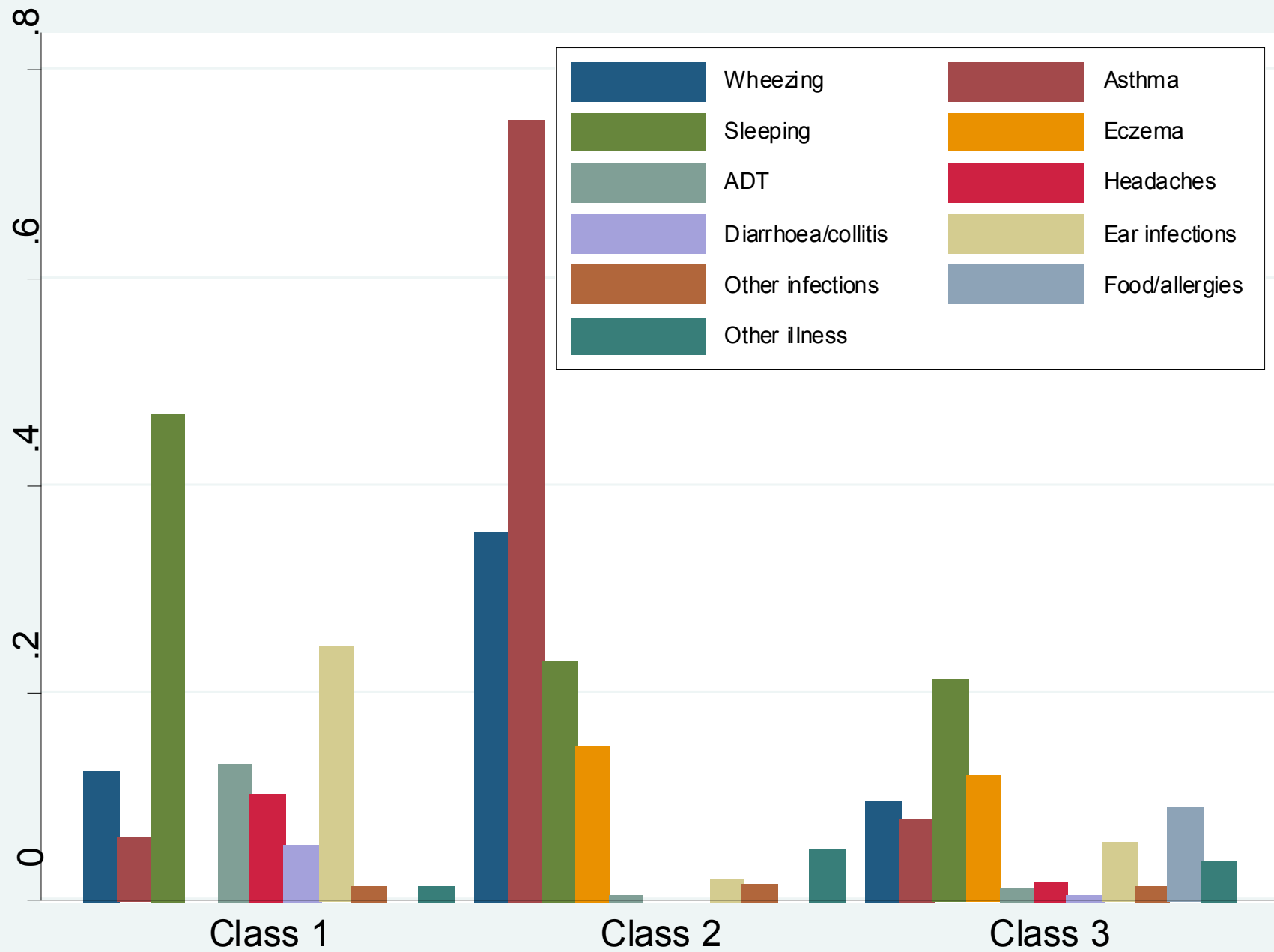


Class 1

Class 2

Latent classes of injuries at w1

	The Latent classes of chronic health conditions at W 2						Naïve sample estimate		Complex sample estimate
	Class 1		Class 2		Class 3				
	n	%	n	%	n	%	n	%	%
<b>6/7 – f2f b8.1 – sc had illness with wheezing</b>	169	12.4%	1044	35.3%	3251	9.4%	4464	15.6%	16.2%
<b>6/7 – f2f b9.1 – sc diagnosed with asthma</b>	169	5.9%	1043	75.1%	3243	7.6%	4455	23.3%	23.7%
<b>6/7 – f2f b14.1 – sc’s sleeping patterns a problem</b>	169	46.7%	1044	22.9%	3248	21.3%	4461	22.6%	23.1%
<b>6/7 – f2f b11.2.1 – eczema</b>	169	0.0%	1044	14.8%	3251	11.8%	4464	12.1%	12.0%
<b>6/7 – f2f b11.2.2 – attention deficit disorder</b>	169	13.0%	1044	0.5%	3251	1.1%	4464	1.4%	1.6%
<b>6/7 – f2f b11.2.3 – frequent headaches</b>	169	10.1%	1044	0.0%	3251	1.8%	4464	1.7%	1.7%
<b>6/7 – f2f b11.2.4 – diarrhoea/collitis</b>	169	5.3%	1044	0.0%	3251	0.4%	4464	0.5%	0.5%
<b>6/7 – f2f b11.2.5 – ear infections</b>	169	24.3%	1044	2.0%	3251	5.6%	4464	5.5%	5.7%
<b>6/7 – f2f b11.2.6 – other infections</b>	169	1.2%	1044	1.5%	3251	1.3%	4464	1.4%	1.4%
<b>6/7 – f2f b11.2.7 – food or digestive allergies</b>	169	0.0%	1044	0.0%	3251	8.7%	4464	6.4%	6.2%
<b>6/7 – f2f b11.2.9 – other illnesses</b>	169	1.2%	1044	4.9%	3251	3.7%	4464	3.9%	4.0%



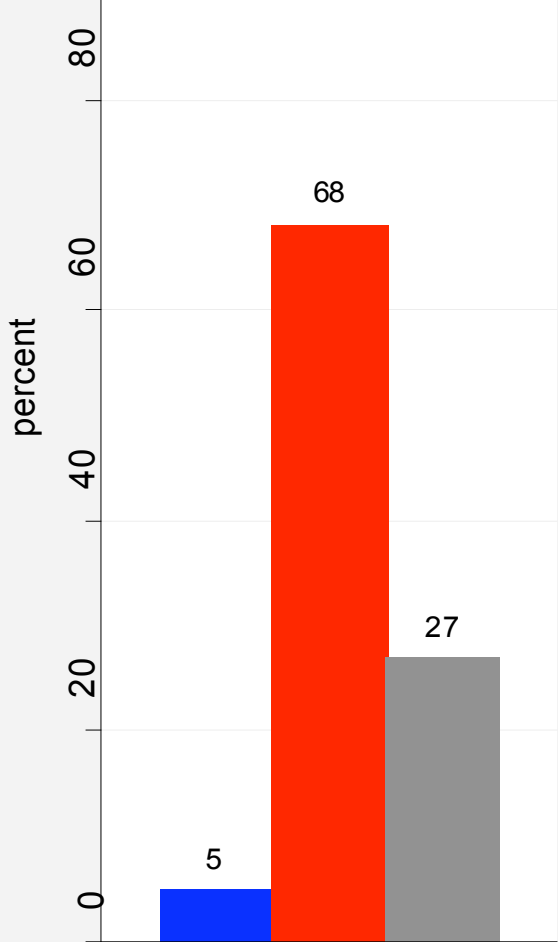
Latent classes of chronic health conditions at Wave 2

	Naive sample estimate		Complex sample
	n	%	%
6/7 – f2f b12.2.1 – injury – broken bones	4464	3.5%	3.5%
6/7 – f2f b12.2.2 – injury – burn or scald	4464	0.4%	0.5%
6/7 – f2f b12.2.3 – injury – dislocation	4464	0.2%	0.2%
6/7 – f2f b12.2.4 – injury – sprain or strain	4464	1.9%	1.9%
6/7 – f2f b12.2.5 – injury – cut or scrape	4464	6.3%	6.3%
6/7 – f2f b12.2.6 – injury – concussion	4464	0.9%	1.0%
6/7 – f2f b12.2.7 – injury – internal (not head)	4464	0.2%	0.2%
6/7 – f2f b12.2.8 – injury – dental	4464	0.7%	0.7%
6/7 – f2f b12.2.9 – injury – accidental poisoning	4464	0.0%	0.0%
6/7 – f2f b12.2.10 – injury – other	4464	3.2%	3.5%

- ✦ How can we make sense of these classes ?
- ✦ One way to call them may be as follows
- ✦ Note that different people may call them differently

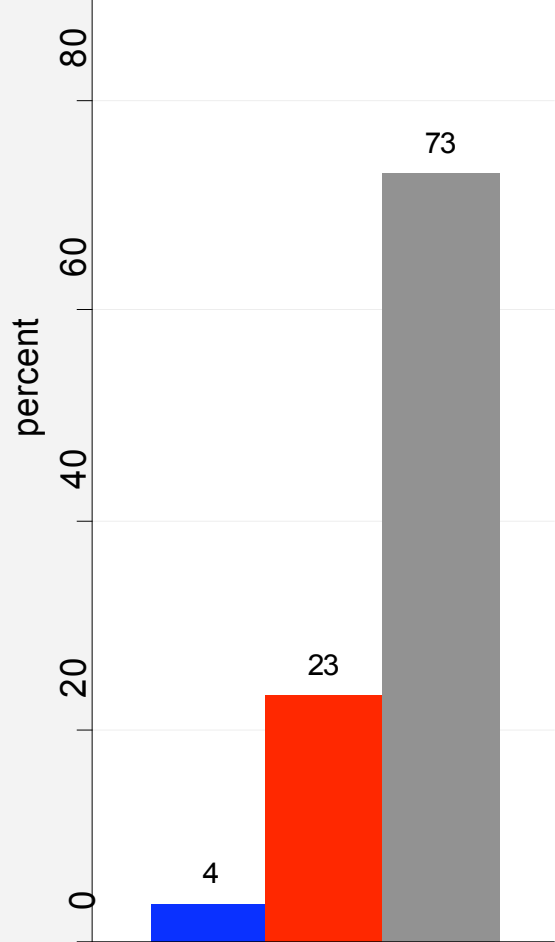


Health conditions: Wave 1



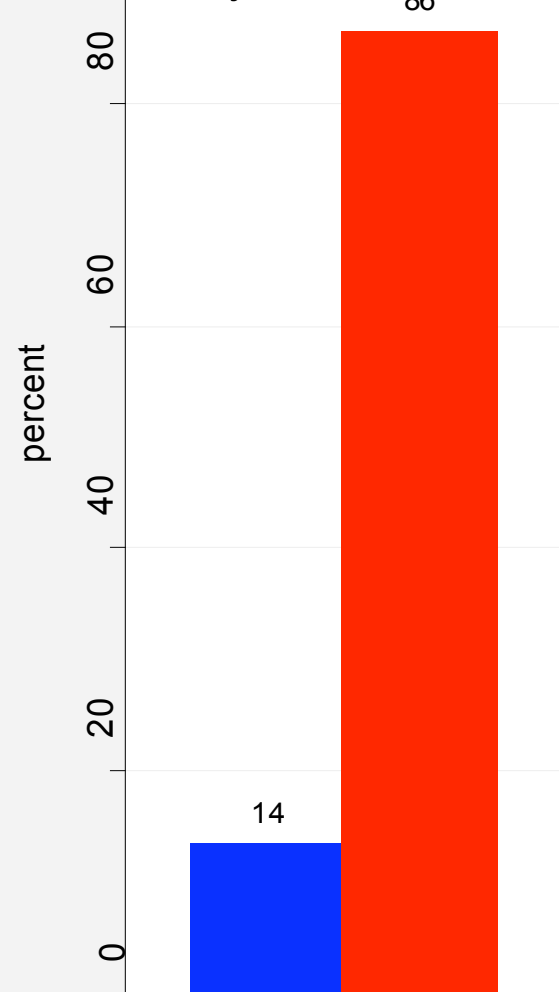
- W1: Worst Health except Wheezing
- W1: Best Group
- W1: Worst resp + fair/bad others

Health conditions: Wave 2



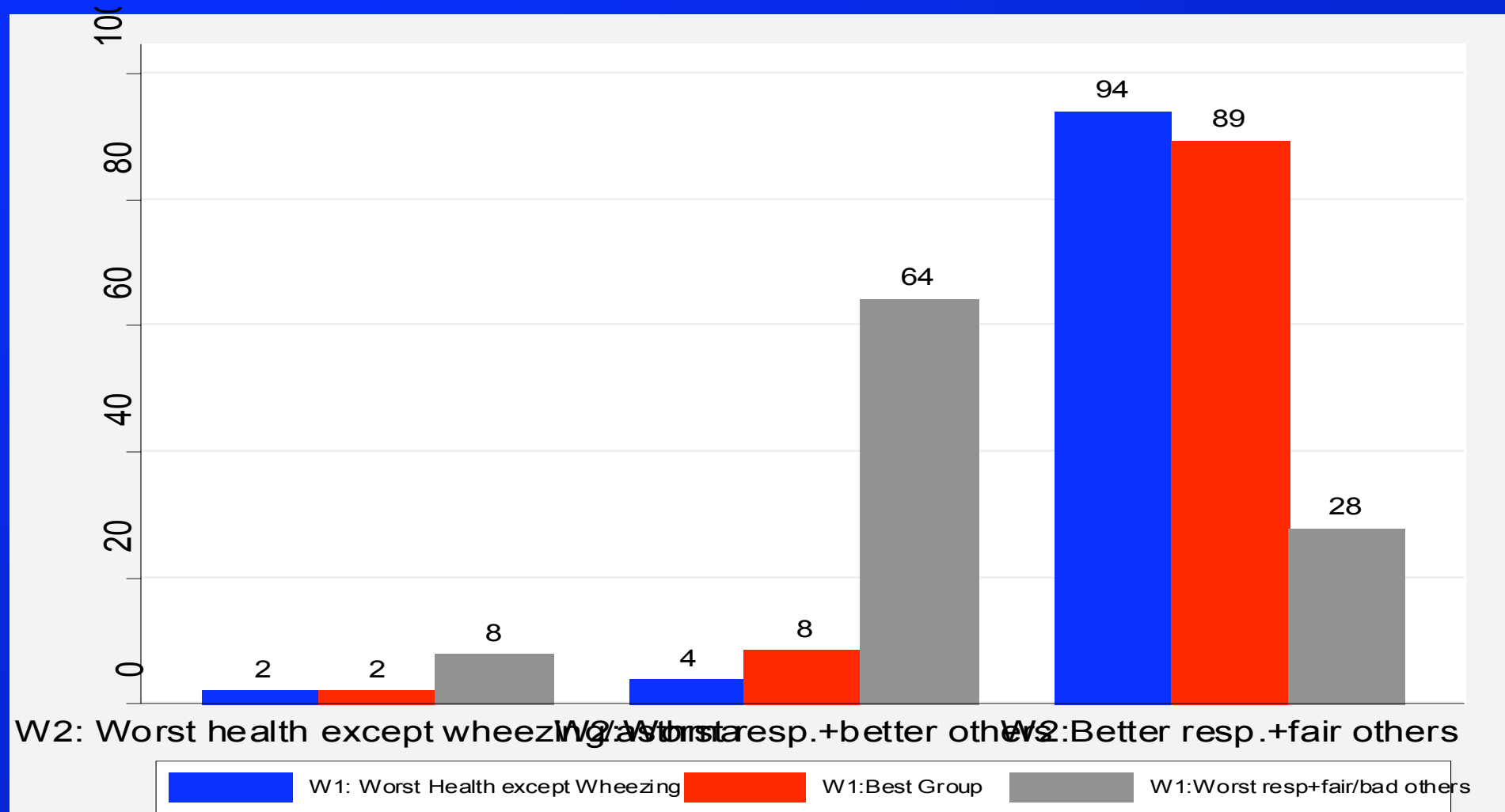
- W2: Worst health except wheezing/asthma
- W2: Worst resp. + better others
- W2: Better resp. + fair others

Injuries: Wave 1

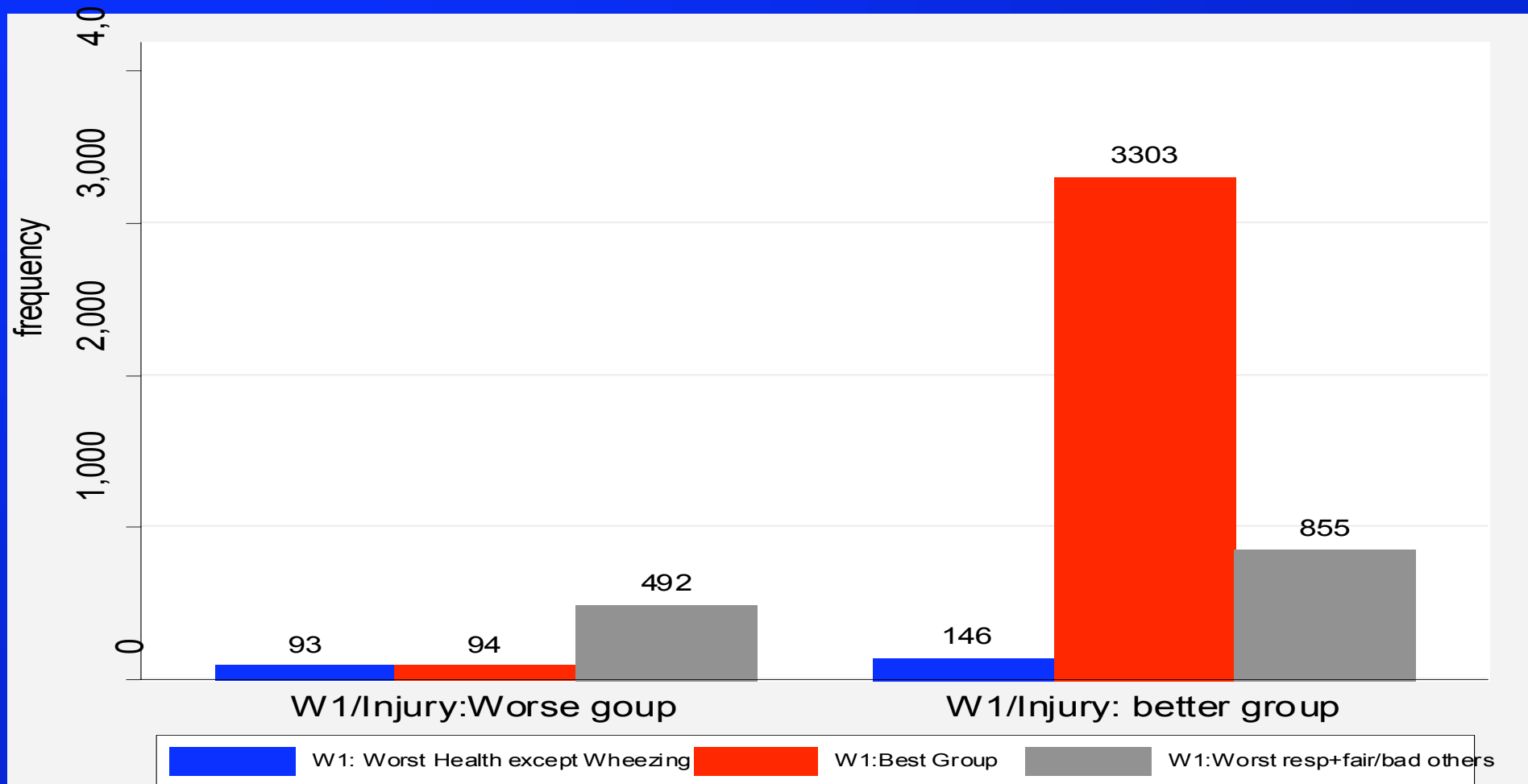


- W1/Injury: Worse group
- W1/Injury: better group

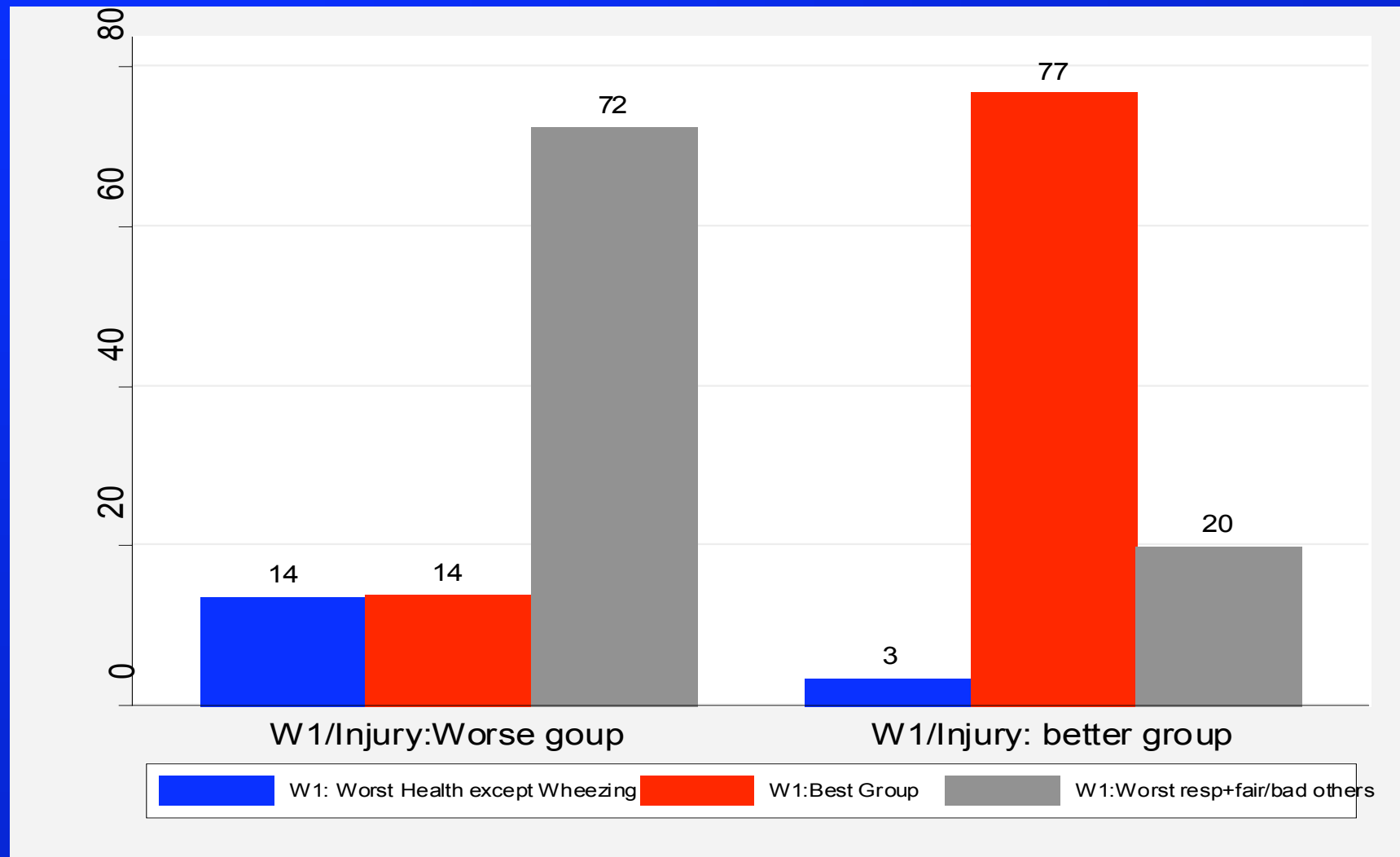
# The transition between baseline and follow-up latent classes - chronic health conditions



# Associations between baseline health conditions and injuries



# Associations between baseline health conditions and injuries



# *What are the predicting factors of these classes*

- ✦ The framework;
- ✦ The MI imputation
- ✦ The results



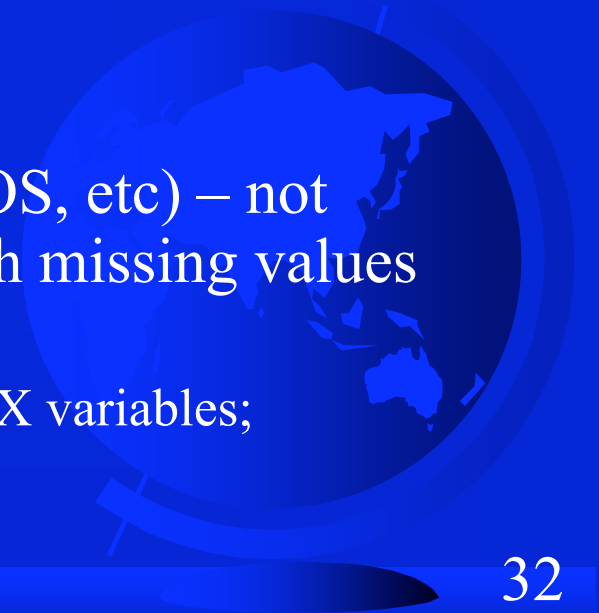
*Methods in dealing with missing values*

- ✦ Littler & Rubin terminology:
- ✦ Missing complete at random (MCAR);
- ✦ Missing at random (MAR);
- ✦ Non-ignorable missing



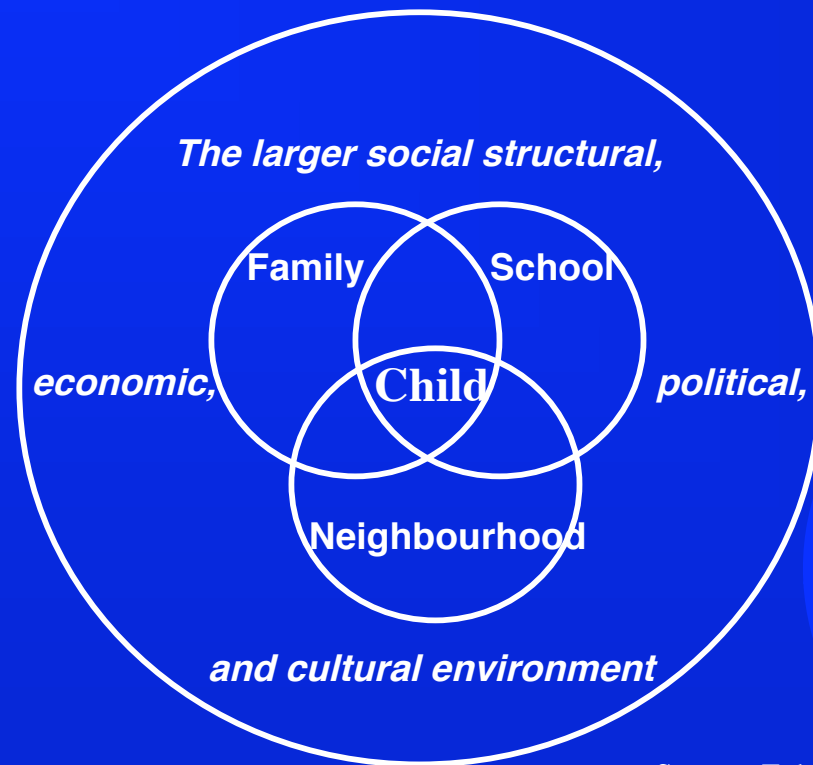
# For MCR

- ✦ Multiple Imputations:
- ✦ SAS Proc MI: joint distribution of all variables
  - Pros: works well with multivariate normal data;
  - Sound theoretical base;
- ✦ Stata ice: imputation with chained equation (ICE);
  - each conditional density of a variable given all other variables
  - Simulation studies worked well;
  - Less strong in distributional assumption;
- ✦ Full maximum likelihood method (MPlus, AMOS, etc) – not imputation, using latent variables in dealing with missing values
  - Pro: simple and intuitively appealing;
  - Cons: cumbersome in dealing with missing values in X variables;
  - Mplus has features in dealing with MI datasets;



# Broad conceptual framework

- ◆ a holistic approach to child development
- ◆ an ecological model (Bronfenbrenner 1979)



Source: Zubrick, Williams & Silburn, 2000

# The ecological predictors

	Sample statistics	
	n	mean or %
4/5 - home - seifa advantage/disadvantage	4983	1004.772
4/5 - home - % with atsi origins for postcode	4983	2.125
4/5 - home - % completed year 12 for postcode	4983	39.763
4/5 - home - % working for postcode	4983	58.720
4/5 - home - % family income <\$1k/week for postcode	4983	52.703
4/5 - home - % speak english for postcode	4983	85.363
4/5 - region of residence (xMET=1)	4983	0.379
4/5 - remoteness area classification	4937	.732
4/5 - fcf 2 - external condition of dwelling	4942	3.649
4/5 - fcf 3 - general condition of nearby buildings	4741	3.701
4/5 - neighbourhood liveability	4976	1.997
4/5 - neighbourhood facilities	4975	1.989
4/5 - socioeconomic position	4965	.109

# The mother's history, family and social variables

4/5 – p1-Currently in the workforce	4972	.574
p1@w1 - f2f a13 - indigenous status	4981	.029
4/5 - p1@w1 - f2f a4 - age	4981	34.744
p1@w1 - f2f a3 – sex (Female=1)	4983	0.971
4/5 - p1 - f2f h5 - highest qualification	3174	3.712
4/5 - p1 - f2f h13 – religion (Yes=1)	4970	.794
4/5 - p1@w1 - f2f a14 - any medical condition/s	4983	1.268
4/5 - f2f k21 - how family is getting on financially	4974	3.201
4/5 - hardship scale	4969	.911
4/5 - f2f k29 - private hospital insurance (yes=1)	4971	.460
p1l d12/p2l c12 - mother drank during pregnancy	4075	.291
p1l d15/p2l c15 - mother smoked during pregnancy	4074	.182
p1l d16/p2l c16 - post-natal depression after pregnancy	3738	.152
4/5 - p1 - frequent binge drinking	4087	.115
4/5 - p1 - p1l d18 – current smoker (yes=1)	4983	.193
4/5 - p1 - k-6 depression scale	4198	4.314
4/5 - p1 - p1l d29 - depression 2 weeks in past year (yes=1)	4180	.319
4/5 - weight status (overweight/obese=1)	4934	0.204
4/5 - parent 1 has a partner	4983	.859
4/5 - stressful life events index	4175	1.635

## The mother's history, family and social variables – continued

4/5 - p1 - home activities index	4980	1.725
4/5 - extra cost activities index	4982	1.039
4/5 - f2f c31 - sc attends school etc.	4982	.954
4/5 - p1 - f2f f1 - parenting self-efficacy rating (1/5: very good/not very good)	4968	2.068
4/5 - p1 - warm parenting	4972	4.441
4/5 - p1 - parent 1 inductive reasoning scale (w1v)	4971	4.263
4/5 - p1 - angry parenting (w1v)	4971	2.179
4/5 - p1 - consistency scale	4969	4.049
4/5 - attachment to family scale	4159	1.899
4/5 - attachment to friends scale	4135	1.845
4/5 - p1 - p1l b19 - level of support by family or friends	3844	1.314
4/5 - p1 - p1l b20 - need support but couldn't get it	2087	2.775
4/5 - neighbourhood belonging scale	4193	2.352
4/5 - p1 - p1l b28 - no. volunteer groups located locally	2856	2.219
4/5 - p1 - coparenting scale	3625	4.428
4/5 - p1 - hendrick relationship quality scale	3621	4.290
Mean of the five items of social contact of P1	4194	4.367

## *Some variables about children*

sc - f2f a3 – sex (female=1)	4983	0.491
4/5 - sc - f2f a12 - main language spoken at home (1=English)	4879	.893
f2f b8 - whether sc ever breastfed	4974	.903
4/5 - no. siblings of sc in household	4983	1.488
Mean of the six items of social contact for SC	4191	3.903
4/5 - sc - f2f b58 - sc's enjoyment of physical activity	4982	4.644
4/5 - serves of fruit and vegetables	4911	2.970
4/5 - serves of high fat food	4933	1.932
4/5 - serves of high fat food (inc. whole milk)	4921	3.489
4/5 - serves of high sugar drinks	4961	1.671

# Advantage of using ICE

- ◆ no multivariate joint distribution assumption; this reason alone makes it appealing since it allows different types of variables to be imputed together;.
- ◆ allowing different kinds of weights, as long as the regression models allow them;
- ◆ easy to understand;
- ◆ easy to use.

# Multiple imputation of 64 variables 1

- ◆ Data saved

- ◆ Further analyzed using micombine in stata



# The predictors of the estimated classes of chronic injuries at baseline

Worse group vs better group

## Predicting variables

	Z-score	P
4/5 - fcf 2 - external condition of dwelling	-1.70	0.089
4/5 - neighbourhood facilities	2.58	0.010
4/5 - p1@w1 - f2f a4 - age	-4.34	0.000
4/5 - p1@w1 - f2f a14 - any medical condition/s	5.04	0.000
4/5 - weight status (overweight/obese=1)	1.72	0.085
4/5 - p1 - home activities index	1.85	0.065
4/5 - extra cost activities index	2.61	0.009
4/5 - p1 - consistency scale	-1.75	0.080
4/5 - attachment to friends scale	-1.65	0.099
4/5 - p1 - p1l b19 - level of support by family or friends	2.07	0.038
sc - f2f a3 – sex (female=1)	-3.23	0.001
4/5 - sc - f2f a12 - main language spoken at home (1=English)	3.14	0.002
4/5 - serves of high fat food	2.61	0.009
4/5 - serves of high fat food (inc. whole milk)	-2.25	0.024

# The predictors of the estimated classes of chronic health condition at baseline

Predicting variables	Worst health vs best health		Worst resp. vs best health	
	Z-score	P	Z-score	P value
4/5 - region of residence (xMET=1)	2.46	0.014	0.35	0.725
4/5 - fcf 3 - general condition of nearby buildings	0.48	0.634	1.71	0.087
4/5 - p1@w1 - f2f a4 - age	-0.63	0.530	-3.19	0.001
4/5 - p1@w1 - f2f a14 - any medical condition/s	2.59	0.010	7.49	0.000
4/5 - hardship scale	0.34	0.730	2.44	0.015
p1l d12/p2l c12 - mother drank during pregnancy	-2.12	0.034	-0.75	0.453
4/5 - weight status (overweight/obese=1)	0.46	0.646	2.49	0.013
4/5 - stressful life events index	0.37	0.708	2.43	0.015
4/5 - p1 - home activities index	3.03	0.002	1.83	0.067
4/5 - extra cost activities index	-0.12	0.905	2.28	0.022
4/5 - p1 - angry parenting (w1v)	1.80	0.072	2.46	0.014
4/5 - p1 - p1l b19 - level of support by family or friends	-0.59	0.552	1.94	0.052
Mean of the five items of social contact of P1	0.29	0.771	1.68	0.093
sc - f2f a3 – sex (female=1)	-1.13	0.258	-3.85	0.000
f2f b8 - whether sc ever breastfed	-0.57	0.568	-2.97	0.003
4/5 - serves of high fat food	0.33	0.744	2.56	0.010
4/5 - serves of high fat food (inc. whole milk)	-0.99	0.321	-2.90	0.004

# The predictors of the estimated classes of chronic health condition at follow-up

Predicting variables	Worst health vs best resp.		Worst resp. vs best resp.	
	Z-score	P	Z-score	P value
4/5 - home - % working for postcode	2.23	0.026	0.94	0.349
4/5 - p1@w1 - f2f a4 - age	0.38	0.707	-2.30	0.022
4/5 - p1 - f2f h13 – religion (Yes=1)	-1.15	0.252	-2.15	0.032
4/5 - p1@w1 - f2f a14 - any medical condition/s	1.50	0.135	4.32	0.000
4/5 - weight status (overweight/obese=1)	1.18	0.237	2.10	0.036
4/5 - parent 1 has a partner	-0.20	0.845	1.83	0.068
4/5 - p1 - angry parenting (w1v)	1.78	0.076	0.26	0.794
4/5 - p1 - p1l b19 - level of support by family or friends	1.41	0.160	1.69	0.090
sc - f2f a3 – sex (female=1)	-1.85	0.065	-3.29	0.001
f2f b8 - whether sc ever breastfed	-1.59	0.112	-1.81	0.070
4/5 - no. siblings of sc in household	0.03	0.974	-2.06	0.039
4/5 - serves of high fat food (inc. whole milk)	0.09	0.932	2.73	0.006
4/5 - serves of high sugar drinks	2.32	0.021	0.76	0.448

# Summary

- ✦ Chronic health conditions have three distinctive groups at both baseline and follow-up; 5% of the all children belongs to the worst category;
- ✦ Injuries have two distinctive groups at baseline but emerged to one homogenous group at follow-up;
- ✦ For each distinctive groups, we identified a group of predicting variables
- ✦ The Latent class analysis may be useful in identifying the important intervention initiatives

Thanks to everyone involved in  
LSAC !!!

What a remarkable study !!!

