



# Classification tree Analysis of Diet, Dental Problems and Obesity in 3 year old children

Crowe M<sup>1</sup>, O'Sullivan, A<sup>2</sup>, Cassetti, O<sup>1</sup>, McGrath C<sup>3</sup>, and  
O'Sullivan M<sup>1</sup>

<sup>1</sup>Dublin Dental University Hospital, Trinity College  
Dublin, <sup>2</sup>Institute Food and Health, University College  
Dublin, <sup>3</sup>Faculty of Dentistry, The University of Hong  
Kong





# Outline

1. Background-common risk factor/data pre-schoolers/Decision Trees
2. Data analysis-modeller/CHAID
3. Results- Ethnicity, Illness, Income, PCG BMI
4. Conclusions and Policy implications
5. Future work



# Jamie's Sugar Rush

A screenshot of a web browser displaying an article page. The browser's address bar shows the URL 'http://www.nature.com/articles/bdjteam2015122'. The page has a purple navigation bar at the top with links for 'Home', 'Current Issue', 'Archive', 'About the magazine', 'About the BDA', 'Contact', 'Digital editions', 'CPD', 'Jobs', 'Advertising &amp; sponsorship', and 'FAQs'. Below the navigation bar, the breadcrumb trail reads 'Home &gt; Features &gt; Article'. The main heading is 'BDJ TEAM | FEATURE' followed by the article title 'Jamie's Sugar Rush – what do we think?'. Below the title, it says 'BDJ Team 2, Article number: 15122 (2015) | doi:10.1038/bdjteam.2015.122' and 'Published online 25 September 2015'. There are buttons for 'PDF', 'Citation', 'Reprints', 'Rights &amp; permissions', and 'Article metrics'. A short text snippet reads: 'It shocked the nation into action, with his petition reaching 100,000 signatures within 48 hours of the documentary airing, but what did we as a profession make of it? After all, the British Dental Journal has been banging on about sugar for 100 years, so what impact could a 60 minute show by a TV chef possibly have? We ask five healthcare professionals for their opinion.' To the right of the article, there is a box titled 'About BDJ Team' with a description of the magazine. Below that are social media links for 'E-alert', 'RSS', 'Facebook', and 'Twitter'. The BDA logo (British Dental Association) is also present. At the bottom right, there is a large white box with purple text that says 'Sign up to the BDJ Team e-alert'. At the bottom of the browser window, a row of colorful circles (red, purple, pink, yellow) is visible.

# Common Risk Factors

## Consumption patterns in Children?

Adverse effects of poor diet:- from “Dental to Mental”



Images courtesy Prof Pat Wall



# Dental Problems



Images courtesy Dr A.O'Connell



# Dental Caries & Overweight Prevalence - Preschool

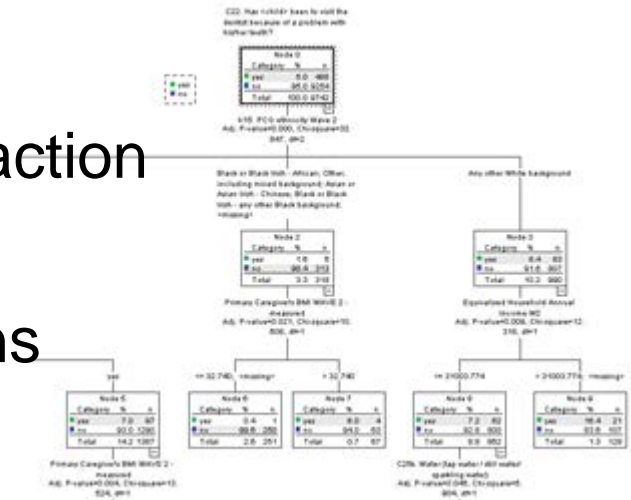
- Increased Prevalence BOTH since 1990's
- EU:Caries: 20-40%: 2-5 year olds
- EU:Obesity/Overweight: 5-10%/15-20%: 4-5 year olds
- IRL:Caries: ???
- IRL:Obesity/Overweight: 3-7%/15-16%: 2-4 year olds





# Decision Trees

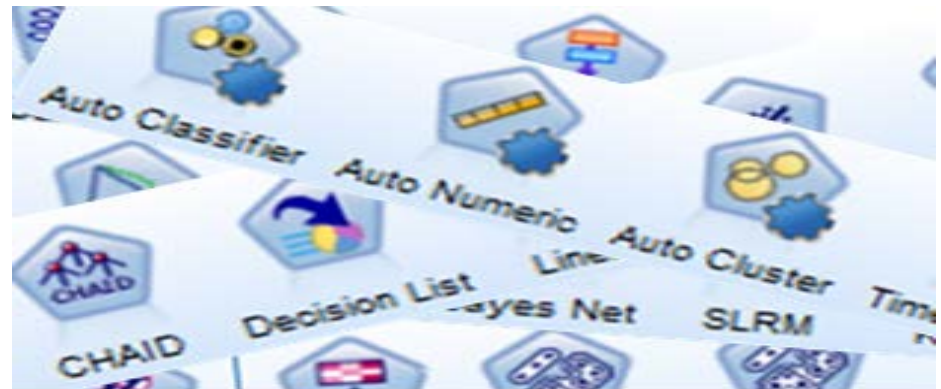
- Tree shaped structures- represent sets of decisions
- Classification- separates data according to outcome (target) variable
- Regression- needed when target is continuous variable
- Recursive partitioning based on interaction
- Visualisation of significant associations





# Terms/Advantages

- CHAID (Chi-square Automatic Interaction Detection)
- Nodes: Root-terminal-leaf
- Mixture of variable types in same analysis
- Detect non-linear interactions
- Not distribution dependant







# Participants

- Data derived from the infant cohort of the *Growing Up in Ireland* (GUI) study.
- Nationally representative sample of 9-month olds in 2007/2008 followed-up at age 3 years in 2010/2011.



# Model Variables

- Target variable = Dental problem
- Physical measures - Height/Weight
- Range of sociodemographic, behavioural, educational and household data measures.
- Child BMI- IOTF classification
- Food Frequency Questionnaire
- Toothbrushing, soothers, accidents, TV viewing
- Reweighted data



# Data analysis - SPSS Modeler

The screenshot displays the IBM SPSS Modeler software interface. The main workspace shows a workflow diagram starting with a data source "XGUI Data\_InfantCoho..". This data is processed by a "Types" node, which then branches into several paths: "Reclassify1", "BML\_Class\_3YO\_Obese...", "No Targets", "BML\_Class\_3YO\_Obese...", "bdprn06c", and "bdprn06c". The "bdprn06c" nodes are connected to "CHAID" nodes, which in turn connect to "BML\_Class\_3YO\_Obese..." nodes. The right-hand pane shows the "Streams" tab with "No ADP - ch..." and "After ADP - c...". Below this, the "CRISP-DM" class structure is visible, including "Business Understanding", "Data Understanding", "Data Preparation", "Modeling", "Evaluation", and "Deployment". The bottom toolbar includes various modeling tools such as "Auto Classifier", "Auto Numeric", "Auto Cluster", "Time Series", "C&R Tree", "Quest", "CHAID", "Decision List", "Linear", "Regression", "PCA/Factor", "Neural Net", "C5.0", "Feature Selection", "Discriminant", "Logistic", "GenLin", "Cox", "SVM", "Bayes Net", "SLRM", and "Apriori". The system tray at the bottom shows the date and time as 15/11/2015, 18:27.

# Sample Description

<b>Child</b>		
<b>BMI Categories (IOTF)</b>	<b>%</b>	<b>N</b>
<b>Underweight</b>	5.7	(557)
<b>Normal</b>	68.3	(6685)
<b>Overweight</b>	17.7	(1737)
<b>Obese</b>	5.7	(559)
<b>Missing</b>	2.6	(256)
<b>Dental Problems in last 12 months</b>	5.0	(493)
<b>Longstanding illness or disability</b>	15.8	(1543)
<b>Hospital admission (ever)</b>	16.1	(1569)

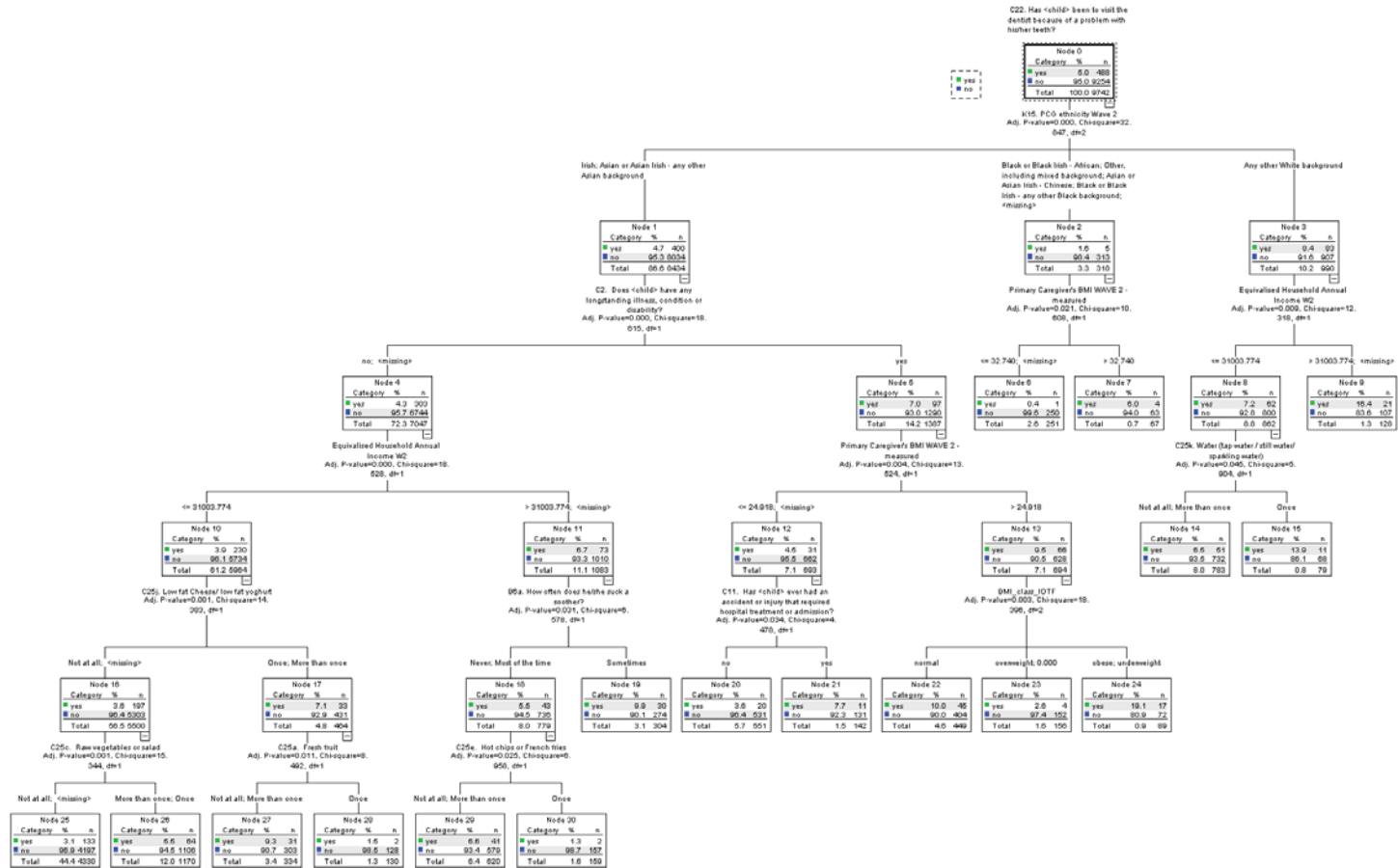
# Sample profile PCG; Mean (SD) or N (%)

PCG	Age (years)	29.6	(6.1)
	BMI (kg/m <sup>2</sup> )	25.99	(5.16)
	Male	26.96	(4.01)
	Female	25.88	(4.91)
	Ethnicity		
	Irish	8261	(84.4)
	White non Irish	1018	(10.4)
	Black	252	(2.6)
	Asian	202	(2.1)
	Other	54	(0.6)
	Equivalised Annual Income	17,874	(9,551)

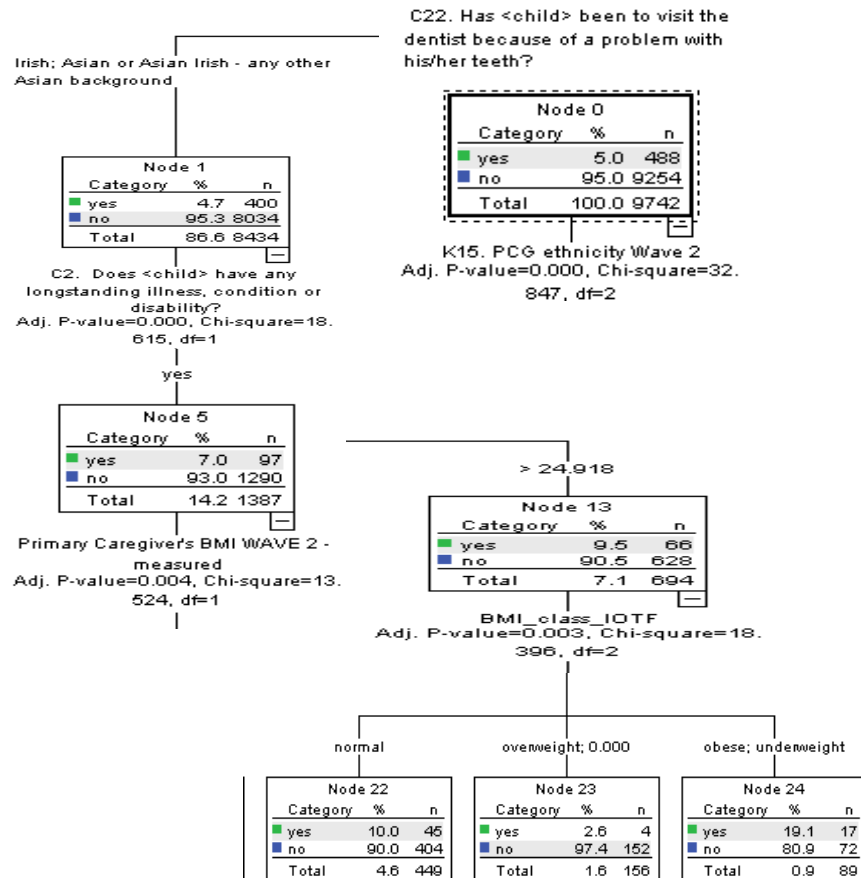




# Classification Tree



# Results 1- Longstanding Illness



Irish; Asian or Asian Irish - any other Asian background

Node 1		
Category	%	n
yes	4.7	400
no	95.3	8034
Total	86.6	8434

C2. Does <child> have any longstanding illness, condition or disability?  
Adj. P-value=0.000, Chi-square=18.615, df=1

yes

Node 5		
Category	%	n
yes	7.0	97
no	93.0	1290
Total	14.2	1387

Primary Caregiver's BMI WAVE 2 - measured  
Adj. P-value=0.004, Chi-square=13.524, df=1

C22. Has <child> been to visit the dentist because of a problem with his/her teeth?

Node 0		
Category	%	n
yes	5.0	488
no	95.0	9254
Total	100.0	9742

K15. PCG ethnicity Wave 2  
Adj. P-value=0.000, Chi-square=32.847, df=2

> 24.918

Node 13		
Category	%	n
yes	9.5	66
no	90.5	628
Total	7.1	694

BMI\_class\_IOTF  
Adj. P-value=0.003, Chi-square=18.396, df=2

normal

overweight; 0.000

obese; underweight

Node 22		
Category	%	n
yes	10.0	45
no	90.0	404
Total	4.6	449

Node 23		
Category	%	n
yes	2.6	4
no	97.4	152
Total	1.6	156

Node 24		
Category	%	n
yes	19.1	17
no	80.9	72
Total	0.9	89

C22. Has <child> been to visit the dentist because of a problem with his/her teeth?

Node 0		
Category	%	n
yes	5.0	488
no	95.0	9254
Total	100.0	9742

K15. PCG ethnicity Wave 2  
Adj. P-value=0.000, Chi-square=32.047, df=2

yes  
no

h; Asian or Asian Irish - any other Asian background

Node 1		
Category	%	n
yes	4.7	400
no	95.3	8034
Total	86.6	8434

C2. Does <child> have any longstanding illness, condition or disability?  
l; P-value=0.000, Chi-square=18.615, df=1

Black or Black Irish - African; Other, including mixed background; Asian or Asian Irish - Chinese; Black or Black Irish - any other Black background; <missing>

Node 2		
Category	%	n
yes	1.6	5
no	98.4	313
Total	3.3	318

Primary Caregiver's BMI WAVE 2 - measured  
Adj. P-value=0.021, Chi-square=10.608, df=1

Any other White background

Node 3		
Category	%	n
yes	8.4	83
no	91.6	907
Total	10.2	990

Equalized Household Annual Income W2  
Adj. P-value=0.009, Chi-square=12.318, df=1

yes

Node 5		
Category	%	n
yes	7.0	97
no	93.0	1290
Total	14.2	1387

Primary Caregiver's BMI WAVE 2 - measured  
Adj. P-value=0.004, Chi-square=13.524, df=1

<= 32,740; <missing>

Node 6		
Category	%	n
yes	0.4	1
no	99.6	250
Total	2.6	251

> 32,740

Node 7		
Category	%	n
yes	6.0	4
no	94.0	63
Total	0.7	67

<= 31003.774

Node 8		
Category	%	n
yes	7.2	62
no	92.8	800
Total	8.8	862

> 31003.774; <missing>

Node 9		
Category	%	n
yes	16.4	21
no	83.6	107
Total	1.3	128

C25k. Water (tap water / still water / sparkling water)  
Adj. P-value=0.045, Chi-square=5.904, df=1

Not at all; More than once

Node 14		
Category	%	n
yes	6.5	51
no	93.5	732
Total	8.0	783

Once

Node 15		
Category	%	n
yes	13.0	11
no	86.1	68
Total	0.8	79

missing

n
73
1010
1083

she suck a  
i-square=6.

<= 24,918; <missing>

Node 12		
Category	%	n
yes	4.5	31
no	95.5	662
Total	7.1	693

C11. Has <child> ever had an accident or injury that required hospital treatment or admission?  
Adj. P-value=0.034, Chi-square=4.478, df=1

> 24,918

Node 13		
Category	%	n
yes	9.5	66
no	90.5	628
Total	7.1	694

BMI\_class\_IOTF  
Adj. P-value=0.003, Chi-square=18.396, df=2

Sometimes

Node 19		
Category	%	n
yes	9.9	30
no	90.1	274
Total	3.1	304

no

Node 20		
Category	%	n
yes	3.6	20
no	96.4	531
Total	6.7	551

yes

Node 21		
Category	%	n
yes	7.7	11
no	92.3	131
Total	1.5	142

normal

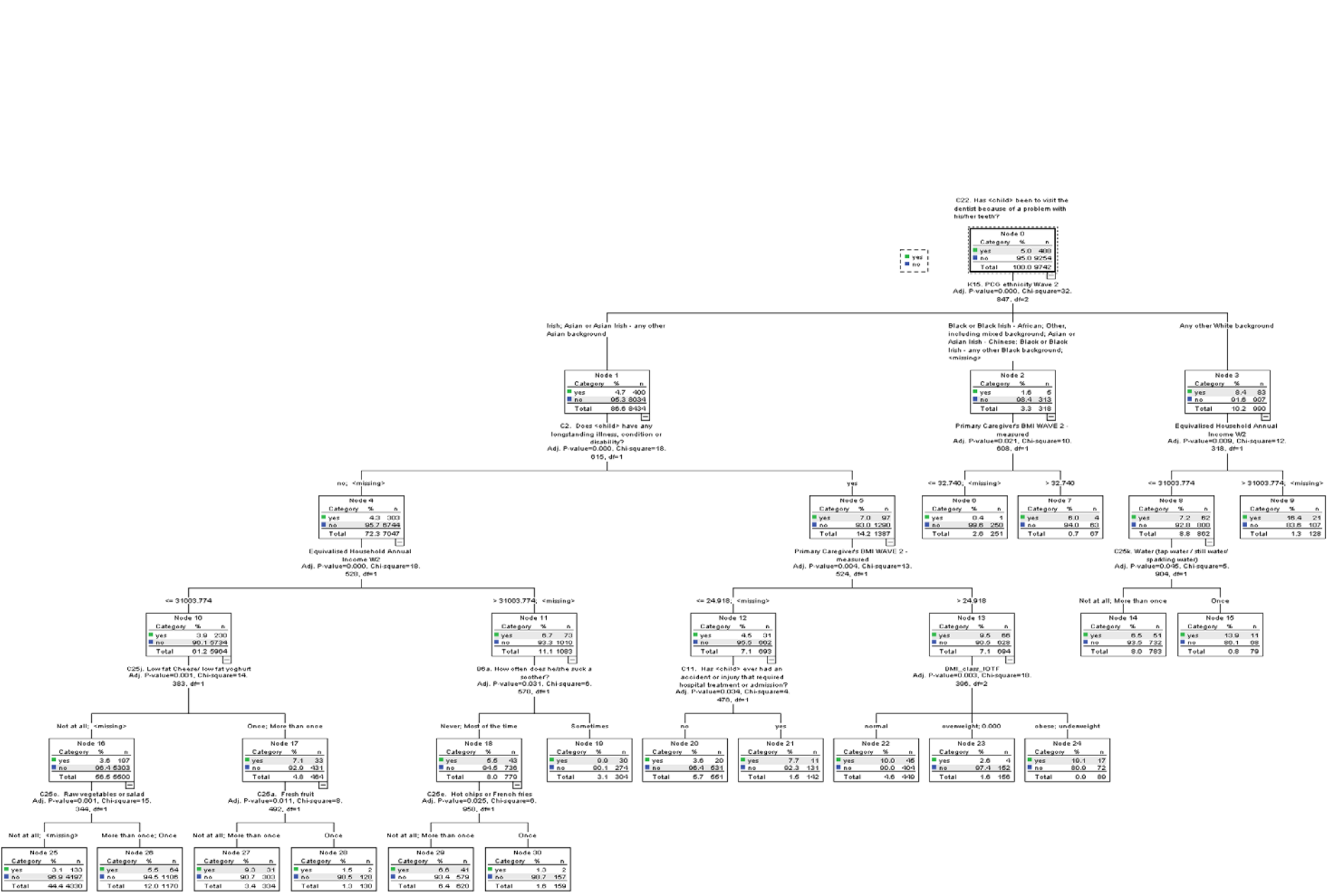
Node 22		
Category	%	n
yes	10.0	46
no	90.0	404
Total	4.6	449

overweight; 0.000

Node 23		
Category	%	n
yes	2.6	4
no	97.4	152
Total	1.6	156

obese; underweight

Node 24		
Category	%	n
yes	10.1	17
no	89.9	72
Total	0.9	89

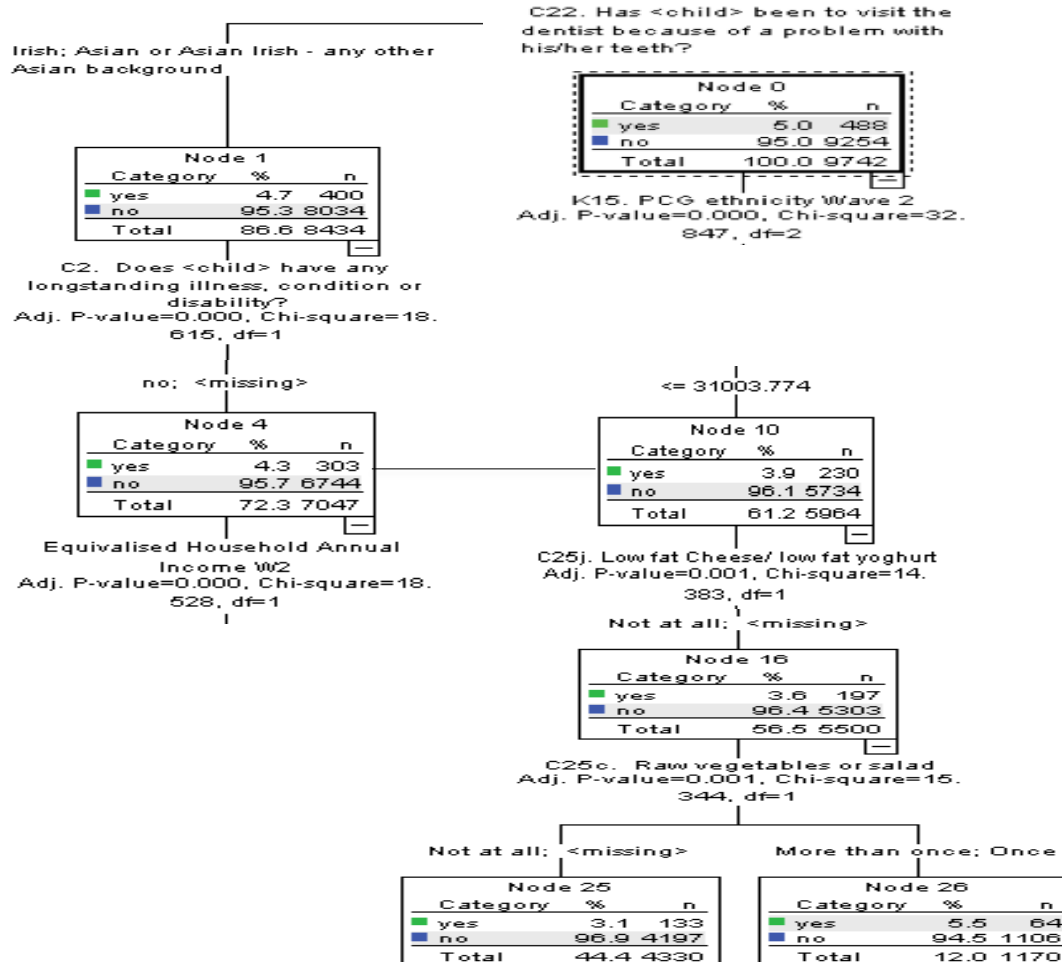






# Results- 2

## No Longstanding Illness



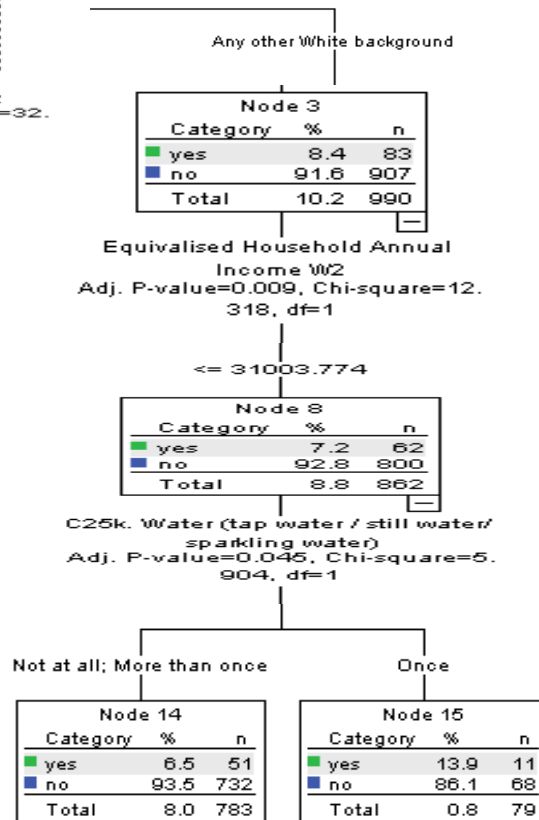


# Results 3- “Other White” ethnicity

C22. Has <child> been to visit the dentist because of a problem with his/her teeth?

Node 0		
Category	%	n
yes	5.0	488
no	95.0	9254
<b>Total</b>	<b>100.0</b>	<b>9742</b>

K15. PCG ethnicity Wave 2  
Adj. P-value=0.000, Chi-square=32.847, df=2





# Model Predictors

- Ethnicity most NB predictor of Dental problem
- Highest prev. Dental Problems: Children obese/underweight with longstanding illness and PCG BMI>24.9
- Food: Low fat cheese/yoghurt. Raw veg/salad, Fresh fruit, French fries - levels 3 and 4 predictors
- Sociodemographic: HH Annual Income, ethnicity
- Oral habits: Soother



# Classification

Observed	Predicted		Percent Correct
	yes	no	
yes	326	162	66.8%
no	3839	5415	58.5%
Overall Percentage	42.8%	57.2%	58.9%

Growing Method: CHAID

Dependent Variable: C22. Has <child> been to visit the dentist because of a problem with his/her teeth?



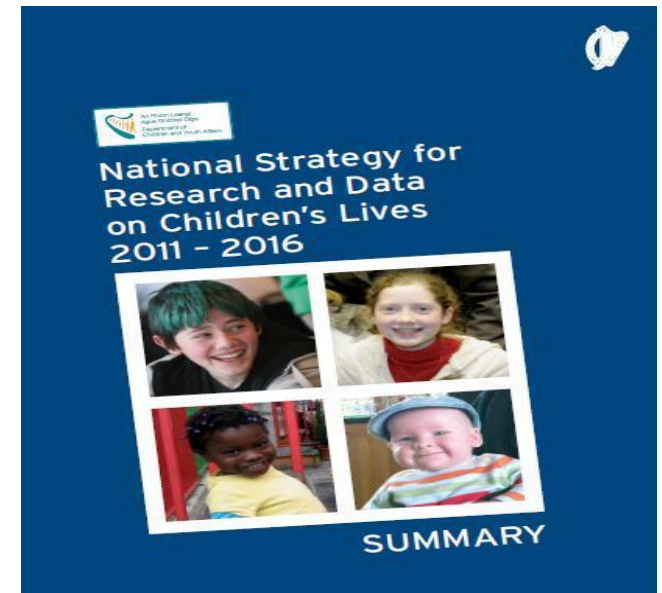
# Conclusions

- Classification trees useful - large survey data
- Complex multilevel variable relationships
- Target subgroups of population cohort
- Disease prevalence data often imbalanced
- Ethnicity most NB predictor
- Food variables- predictors at higher levels
- Obese/underweight AND dental problems





# Policy implications





# Future work

- Dietary pattern using NPNS (IUNA) data
- Parallel Coordinates/data visualisation
- 5 Year old Dataset
- Predictive model



# Acknowledgments

Thanks to:

GUI infants and parents

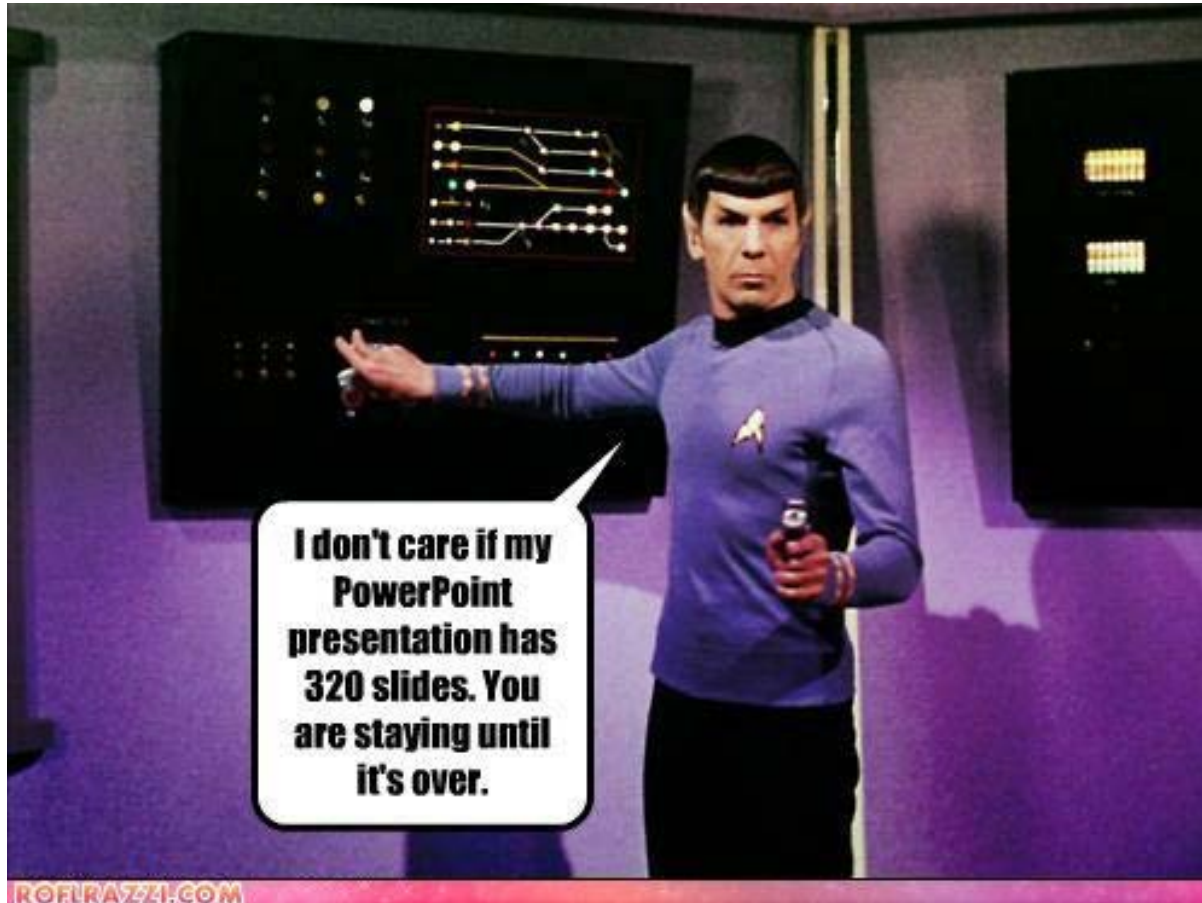
ESRI/GUI team

DDUH



**Growing Up  
in Ireland**  
National Longitudinal  
Study of Children

# Questions?





# References

- Arora, A., Scott, J. A., Bhole, S., Do, L., Schwarz, E., & Blinkhorn, A. S. (2011). Early childhood feeding practices and dental caries in preschool children: a multi-centre birth cohort study. *BMC Public Health*, 11(1), 28.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*: CRC press.
- Hayden, C., Bowler, J. O., Chambers, S., Freeman, R., Humphris, G., Richards, D., & Cecil, J. E. (2012). Obesity and dental caries in children: a systematic review and meta-analysis. *Community Dent Oral Epidemiol*, 41(4), 289-308. doi: 10.1111/cdoe.12014
- Hong, L., Ahmed, A., McCunniff, M., Overman, P., & Mathew, M. (2008). Obesity and Dental Caries in Children Aged 2-6 Years in the United States: National Health and Nutrition Examination Survey 1999-2002. *J Public Health Dent*, 68(4), 227-233. doi: 10.1111/j.1752-7325.2008.00083.x
- Hooley, M., Skouteris, H., Boganin, C., Satur, J., & Kilpatrick, N. (2012). Body mass index and dental caries in children and adolescents: a systematic review of literature published 2004 to 2011. *Syst Rev*, 1(1), 57.
- Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nat Biotechnol*, 26(9), 1011-1013. doi: 10.1038/nbt0908-1011
- Loh, W.-Y. (2014). Fifty Years of Classification and Regression Trees. *International Statistical Review*, 82(3), 329-348. doi: 10.1111/insr.12016
- Meyer, B. D., & Lee, J. Y. (2015). The Confluence of Sugar, Dental Caries, and Health Policy. *J Dent Res*, 94(10), 1338-1340. doi: 10.1177/0022034515598958
- Norberg, C., Hallstrom Stalin, U., Matsson, L., Thorngren-Jerneck, K., & Klingberg, G. (2012). Body mass index (BMI) and dental caries in 5-year-old children from southern Sweden. *Community Dent Oral Epidemiol*, 40(4), 315-322. doi: 10.1111/j.1600-0528.2012.00686.x
- Qadri, G., Alkilzy, M., Feng, Y. S., & Splieth, C. (2015). Overweight and dental caries: the association among German children. *International Journal of Paediatric Dentistry*, 25(3), 174-182.
- Sheiham, A., & James, W. P. (2015). Diet and Dental Caries: The Pivotal Role of Free Sugars Reemphasized. *J Dent Res*, 94(10), 1341-1347. doi: 10.1177/0022034515590377
- Watt, R. G., & Sheiham, A. (2012). Integrating the common risk factor approach into a social determinants framework. *Community Dent Oral Epidemiol*, 40(4), 289-296.
- Whelton, H., O'Mullane, D., Harding, M., Guiney, H., Cronin, M., Flannery, E., & Kelleher, V. (2006). North South survey of children's oral health in Ireland 2002.