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

Crowe M\textsuperscript{1}, O’Sullivan, A\textsuperscript{2}, Cassetti, O\textsuperscript{1}, McGrath C\textsuperscript{3}, and O’Sullivan M\textsuperscript{1}

\textsuperscript{1}Dublin Dental University Hospital, Trinity College Dublin, \textsuperscript{2}Institute Food and Health, University College Dublin, \textsuperscript{3}Faculty of Dentistry, The University of Hong Kong

www.growingup.ie
1. Background-common risk factor/data preschoolers/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 – what do we think?

*BDJ Team* | FEATURE

**Jamie’s Sugar Rush**

*BDJ Team 2, Article number: 15122 (2015) | doi:10.1038/bdjteam.2015.122
Published online 25 September 2015*

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.
Common Risk Factors

Consumption patterns in Children?

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

Images courtesy Prof Pat Wall
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.

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
<table>
<thead>
<tr>
<th>Child</th>
<th>%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI Categories (IOTF)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>5.7</td>
<td>(557)</td>
</tr>
<tr>
<td>Normal</td>
<td>68.3</td>
<td>(6685)</td>
</tr>
<tr>
<td>Overweight</td>
<td>17.7</td>
<td>(1737)</td>
</tr>
<tr>
<td>Obese</td>
<td>5.7</td>
<td>(559)</td>
</tr>
<tr>
<td>Missing</td>
<td>2.6</td>
<td>(256)</td>
</tr>
<tr>
<td>Dental Problems in last 12 months</td>
<td>5.0</td>
<td>(493)</td>
</tr>
<tr>
<td>Longstanding illness or disability</td>
<td>15.8</td>
<td>(1543)</td>
</tr>
<tr>
<td>Hospital admission (ever)</td>
<td>16.1</td>
<td>(1569)</td>
</tr>
<tr>
<td>PCG</td>
<td>Age (years)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>29.6</td>
<td>(6.1)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.99</td>
<td>(5.16)</td>
</tr>
<tr>
<td>Male</td>
<td>26.96</td>
<td>(4.01)</td>
</tr>
<tr>
<td>Female</td>
<td>25.88</td>
<td>(4.91)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish</td>
<td>8261</td>
<td>(84.4)</td>
</tr>
<tr>
<td>White non Irish</td>
<td>1018</td>
<td>(10.4)</td>
</tr>
<tr>
<td>Black</td>
<td>252</td>
<td>(2.6)</td>
</tr>
<tr>
<td>Asian</td>
<td>202</td>
<td>(2.1)</td>
</tr>
<tr>
<td>Other</td>
<td>54</td>
<td>(0.6)</td>
</tr>
<tr>
<td>Equivalised Annual Income</td>
<td>17,874</td>
<td>(9,551)</td>
</tr>
</tbody>
</table>
Results 1-
Longstanding Illness

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

Node 0
Category | % | n
-------|---|---
yes | 5.0 | 438
no | 95.0 | 9254
Total | 100.0 | 9692

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

Node 1
Category | % | n
-------|---|---
yes | 4.7 | 493
no | 95.0 | 8884
Total | 99.7 | 9377

Node 5
Primary Caregiver BMI Wave 2 - measured
Adj. P-value=0.004, Chi-square=13.524, df=1

> 24.918

Node 13
BMI_class_IOTF
Adj. P-value=0.003, Chi-square=18.396, df=2

Node 22
Category | % | n
-------|---|---
yes | 10.0 | 56
no | 90.0 | 494
Total | 100.0 | 550

Node 23
Category | % | n
-------|---|---
yes | 26.0 | 4
no | 74.0 | 152
Total | 100.0 | 156

Node 24
Category | % | n
-------|---|---
yes | 10.0 | 17
no | 90.0 | 72
Total | 100.0 | 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=32.847, df=2

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

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

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
Results - 2
No Longstanding Illness

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

Node 0

<table>
<thead>
<tr>
<th>Category</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>5.0</td>
<td>489</td>
</tr>
<tr>
<td>no</td>
<td>95.0</td>
<td>9250</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>9742</td>
</tr>
</tbody>
</table>

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

C23. Does <child> have any long-standing illness, condition or disability?
Adj. P-value = 0.000, Chi-square = 18.615, df = 1

Node 1

<table>
<thead>
<tr>
<th>Category</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>4.7</td>
<td>403</td>
</tr>
<tr>
<td>no</td>
<td>95.3</td>
<td>8334</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>8737</td>
</tr>
</tbody>
</table>

Equivalized Household Annual Income W2
Adj. P-value = 0.000, Chi-square = 18.528, df = 1

Node 4

<table>
<thead>
<tr>
<th>Category</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>85.7</td>
<td>7047</td>
</tr>
<tr>
<td>no</td>
<td>14.3</td>
<td>1280</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>8327</td>
</tr>
</tbody>
</table>

<= 31000.774

Node 10

<table>
<thead>
<tr>
<th>Category</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>0.0</td>
<td>230</td>
</tr>
<tr>
<td>no</td>
<td>99.9</td>
<td>1730</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>1960</td>
</tr>
</tbody>
</table>

C25j. Low fat cheese/ low fat yoghurt
Adj. P-value = 0.001, Chi-square = 14.993, df = 1

Node 16

<table>
<thead>
<tr>
<th>Category</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>5.0</td>
<td>197</td>
</tr>
<tr>
<td>no</td>
<td>95.0</td>
<td>6303</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>6500</td>
</tr>
</tbody>
</table>

C25c. Raw vegetables or salad
Adj. P-value = 0.001, Chi-square = 15.304, df = 1

Node 25

<table>
<thead>
<tr>
<th>Category</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>5.1</td>
<td>133</td>
</tr>
<tr>
<td>no</td>
<td>94.9</td>
<td>4197</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>4330</td>
</tr>
</tbody>
</table>

Node 26

<table>
<thead>
<tr>
<th>Category</th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>0.6</td>
<td>64</td>
</tr>
<tr>
<td>no</td>
<td>99.4</td>
<td>1117</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>1180</td>
</tr>
</tbody>
</table>
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 | 466
no | 95.0 | 2,054
Total | 100.0 | 2,520

K16. PC0 ethnicity Wave 2
Adj. P-value=0.000, Chi-square=32.847, df=2

Any other White background

Node 5
Category | %  | n
---|---|---
yes | 84.4 | 93
no | 15.6 | 1007
Total | 102 | 900

Equivalised Household Annual Income Wave 2
Adj. P-value=0.000, Chi-square=12.316, df=1

<= 31003.774

Node 8
Category | %  | n
---|---|---
yes | 7.2 | 52
no | 92.8 | 800
Total | 98 | 802

C25k. Water (tap water/ still water/ sparkling water)
Adj. P-value=0.003, Chi-square=6.904, df=1

Not at all; More than once; Once

Node 14
Category | %  | n
---|---|---
yes | 5.5 | 51
no | 93.5 | 732
Total | 89 | 783

Node 15
Category | %  | n
---|---|---
yes | 13.9 | 11
no | 86.1 | 58
Total | 100 | 79
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

<table>
<thead>
<tr>
<th>Observed</th>
<th>Predicted</th>
<th>Percent Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
<td>66.8%</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>58.5%</td>
</tr>
<tr>
<td>Overall Percentage</td>
<td>42.8%</td>
<td>57.2%</td>
</tr>
</tbody>
</table>

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
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
I don't care if my PowerPoint presentation has 320 slides. You are staying until it's over.
References