

# Developing an early warning system combined with dynamic LMS data

Danny Green and Le Van Anh  
RMIT University, Vietnam

# Outline

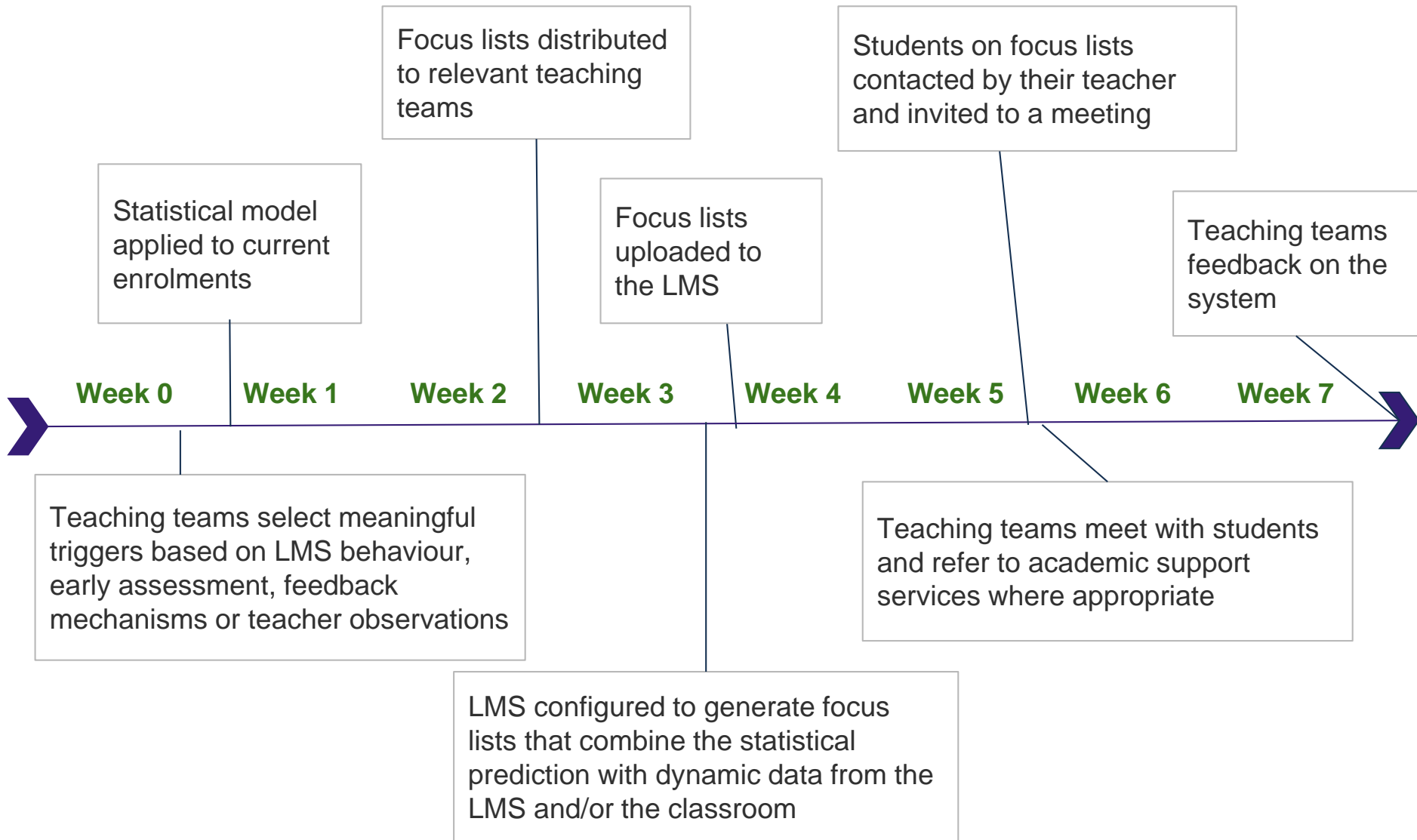
- Background and Rationale
- The EWS
- The General Statistical Model
- The Pilot
- Combining dynamic LMS Data
- All first year courses
- Conclusions
- Future Research

# Background and Rationale

- RMIT Vietnam is an unique context
- Lack of institutional research in this area
- Existing at-risk system has drawbacks
  - Based on fails vs course load
  - Referrals to support services happening too late in the semester
  - Doesn't involve the teacher

# What we wanted to know

- What factors influenced achievement?
- How accurate could the statistics be?
  - What % of low scores could we predict?
- What impact would the model have on achievement?



# The General Model

- $n = 78\ 000$
- 20 variables
- Multiple linear regression and Bayesian Model Averaging (BMA)
- R-squared of 0.6815
- 88% of the students who achieved low scores ( $\leq 55$ ) were predicted by the model's "focus-list".
- Key factors related to past academic performance

# The Pilot

- Three first-year, high enrolment, core business courses ( n =1009)
- Researchers then refined the model to be specifically relevant to the chosen courses
- A training dataset was then created (ratio = 7:3) and BMA was applied to ascertain the most significant predictors of final grades
- Multiple Linear regression was then applied using these predictors to develop a predictive model.
- multiple R-squared of 0.416 (r ~ 0.65)

# Pilot Model

- Volunteers from the teaching teams - 10 teachers
- requested more detailed feedback from teachers
- Researched internal triggers and consulted to find good intervention moments
- Used smart lists in blackboard
- Composed template emails - designed as positive nudges (including consultation with health and well being)



# LMS Triggers

Course		Trigger 1	Trigger 2	Trigger 3	Trigger 4
Introduction to Management	Trigger type	Quiz Attempts	Online Test	Individual Paper	Online test + Individual Paper
	Correlation	$r = .329$	$r = .640$	$r = .843$	$r = .935$
	Threshold	<1	<18	<20	<42
	Correlated to final score	>60	>60	>60	>60
Business Computing	Trigger type	Quiz Attempts	Assessment 1.1 + 1.2	Assessment 1.1+1.2+2	Assessment 2
	Correlation	N/A	$r = .653$	$r = .843$	$r = .671$
	Threshold	Less than 2 attempts	<41	<70	<26
	Correlated to final score	N/A	<60	<60	<60
Marketing Principles	Trigger type	BB Activity	Assessment 1	Assessment 2	
	Correlation	$r = .368$	$r = .407$	$r = .558$	
	Threshold	<1hr 70% below the average	$\leq 16$	$\leq 16$	
	Correlated to final score	<61	<60	<60	

# Pilot Model Accuracy

- 63.22% positive predictions across the three courses
- 71% of all scores within a 10% error.
- Average points difference of 8.94
- 49 of the 110 fails were on the lists
- Combined with data coming from the LMS the prediction accuracy increased

# Model and LMS evaluation

<b>BUSM4185</b>	<b>Model</b>	<b>Trig1</b>	<b>Trig2</b>	<b>Model + Trig 1</b>	<b>Model + Trig 1 + Trig2</b>
Sensitivity	38%	55%	13%	78%	20%
Specificity	88%	81%	94%	65%	100%
<b>Positive Predictive Value</b>	37%	34%	29%	56%	100%
Negative Predictive Value	89%	91%	86%	83%	82%
<b>ISYS2109</b>	<b>Model</b>	<b>Trig1</b>	<b>Trig2</b>	<b>Model + Trig 1</b>	<b>Model + Trig 1 + Trig2</b>
Sensitivity	36%	30%	62%		54%
Specificity	93%	87%	94%	39%	100%
<b>Positive Predictive Value</b>	72%	59%	86%	78%	100%
Negative Predictive Value	74%	67%	80%	25%	43%
<b>MKTG1205</b>	<b>Model</b>	<b>Trig1</b>	<b>Trig2</b>	<b>Model + Trig 1</b>	<b>Model + Trig2</b>
Sensitivity	36%	Not available	59%	Not Available	36%
Specificity	92%		91%		92%
<b>Positive Predictive Value</b>	69%		80%		69%
Negative Predictive Value	73%		77%		73%

# Teacher feedback

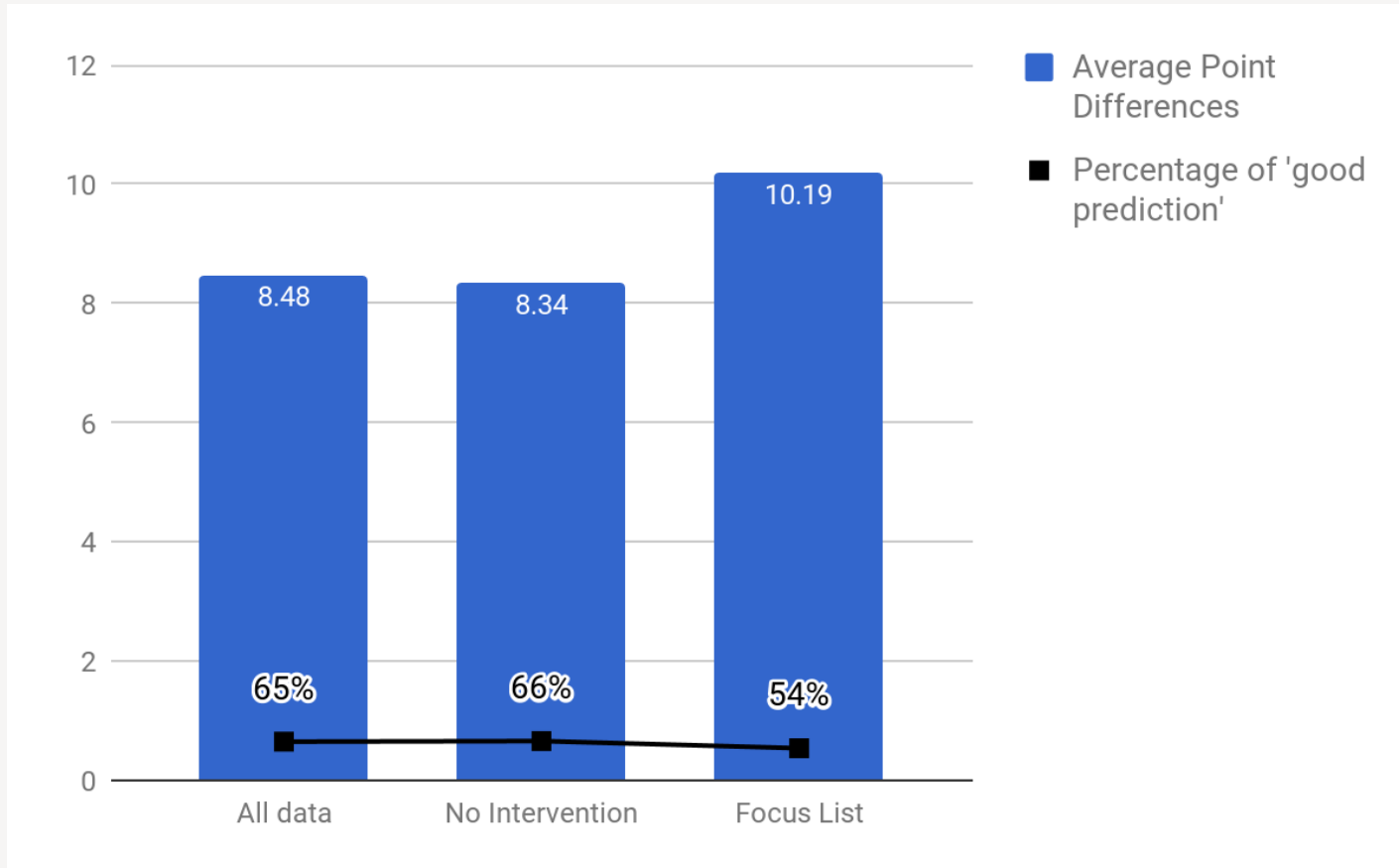
- Teacher feedback uptake:

0%

# All first year courses S1 2017

- n = 4329
- Focus list 547 ( reduced to 462)
- 93 teachers and 48 courses
- Information sessions held with all relevant teaching teams
- Successfully had feedback added as an academic KPI
- Simplified feedback - 3 yes/no drop downs
- Internal Validation = 71% good prediction, 7.77 average points difference, positive predictive value 73.5%

# Actual first year results



# Focus List records

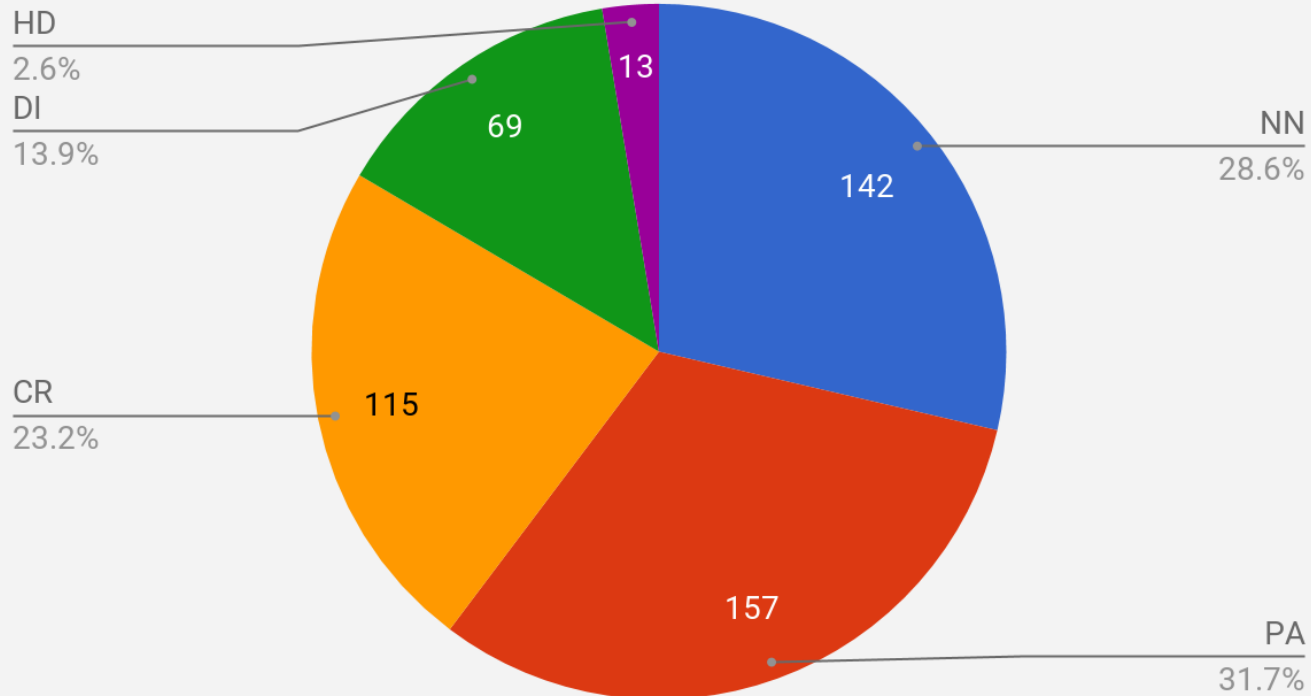
All courses		Actual Score $\leq 55$	$> 55$	Total
Predicted Score	$\leq 55$	238	223	461
	$> 55$	12	23	35
	Total	250	246	496

- 51.63% Positive Predictive Value
- out of the 496 records on the focus lists, 142 did not successfully pass their course
- 97 records had a PS  $\geq 55$  and an AS of  $\geq 55$ , were not contacted, met or referred by the teacher
- Focus list predicted 32% of actual fails across all first year courses

# Focus list records

Results of students the model predicted would achieve  $\leq 55$

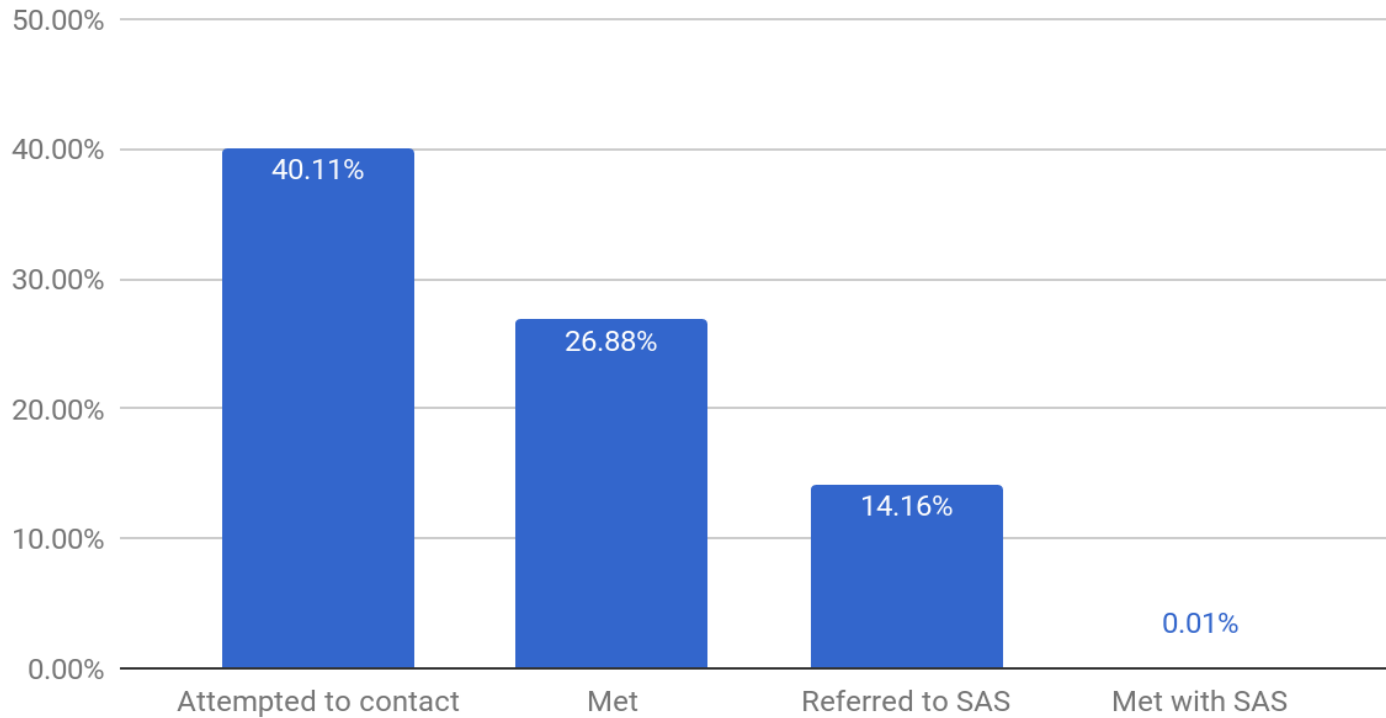
Final grades of focus list records S1 2017





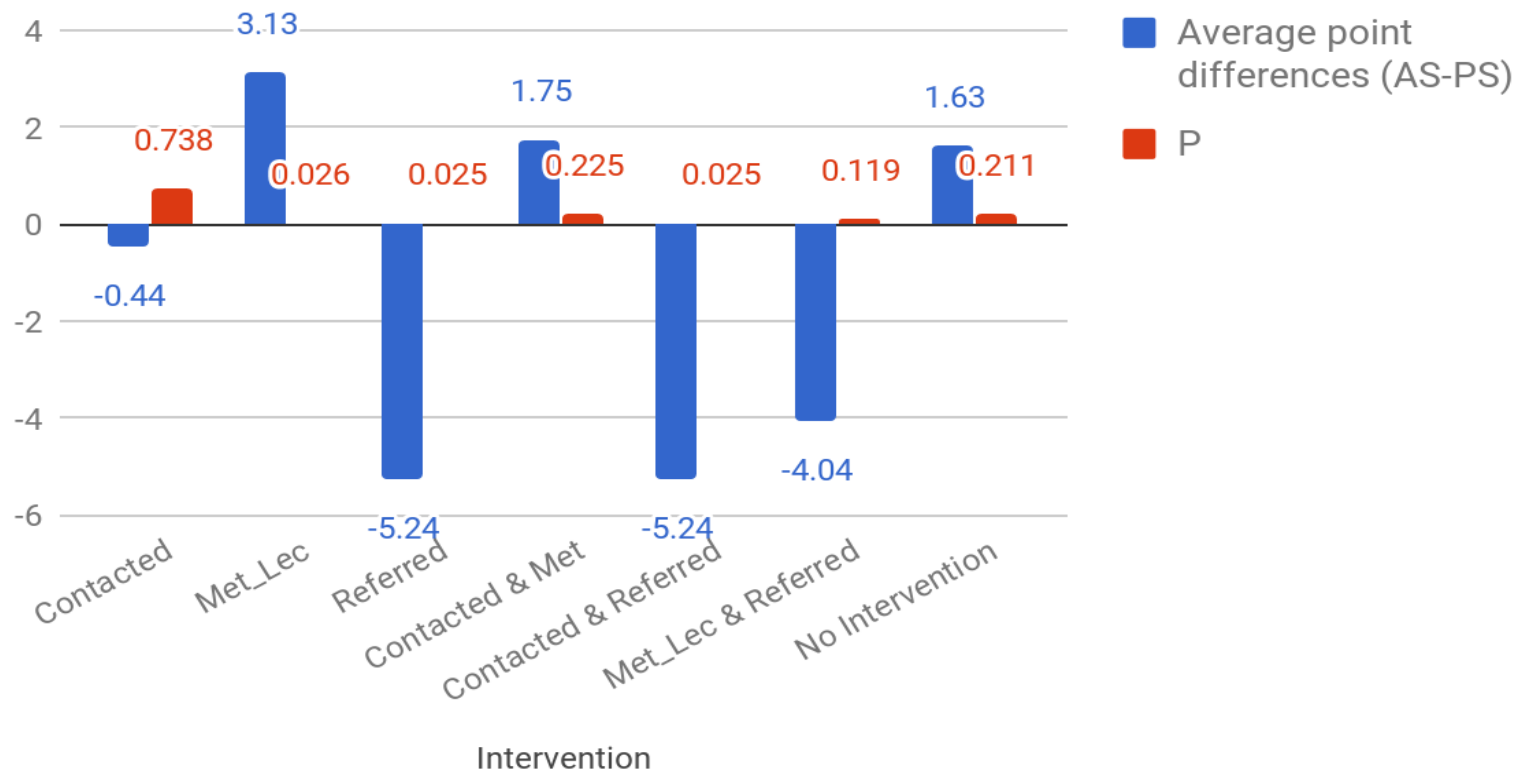
# Interventions

Percentage of Interventions - Focus List Records S1 2017



# Predicted vs Actual results against intervention factors

Average point differences (AS-PS) and P



# Teacher feedback

% of teachers who participated (gave feedback at the end):

**75.27%**

\*but in reality, still have 183 records with no feedback

# Conclusions

- Statistical model already providing a powerful tool for teachers
- It is only a tool, and needs to be combined with in course data to be meaningful
- Further research, modeling and consultation and collaboration with stake holders necessary

# The Future

- Research and Data
  - working with admissions to obtain missing data
  - geographical, family and socioeconomic status data
  - work to achieve greater buy in from faculty
  - machine learning - neural networking
  - LMS behaviour
- Permanent integration into the Student Record System

# Developing an early warning system combined with dynamic LMS data

Danny Green and Le Van Anh  
RMIT University, Vietnam