

Project Title: Usage of Early Alerts Indicators on two level 1 modules

Keywords: Early Alert Indicators, Associate Lecturer views, student views, predictive analytics.

Your name: Carol Calvert

Report submission date: 2nd December 2019

Names of any key staff associated with the project: Carol Calvert (lead), Alison Bromley, Chris Hughes, Clare Morris (AL)

Contact point: Carol Calvert

Summary

The Open University (UK) has developed two systems of predictive modelling to identify students at risk. The main function of one system was to predict whether a student would pass their module whilst the second system was primarily designed to generate weekly predictions on whether a student was likely to submit their next Tutor Marked Assignment. The Early Alerts Indicators combined the two systems within one graphical interface and this interface is now made available to Associate Lecturers and module teams across all undergraduate modules. This eSTEEeM project was designed to establish Associate Lecturers views on how to maximise their use of the current Early Alerts data and to identify improvements based on their recommendations.

The views of the Associate Lecturers', who were recruited as volunteers to the project, were generally positive about the Early Alerts Indicator information. The ALs identified several areas for improvement and amongst these were two key areas: around timing and around the transparency / simplicity of how the probabilities of passing the module are generated.

Whilst the indicators can be made available earlier this comes at a cost in accuracy. Pre module start probabilities have an accuracy of 73%-75% and accuracy increases to around 85% as soon as information on the first assignment is available. The accuracy of predictions on whether a student will pass are in the high 90%'s shortly before the module ends. Simplifying the variables used in the prediction models involves only around 3%-5% percentage points loss in accuracy. Thus, pre-start probabilities of success can be generated, with an accuracy of around 70%, by considering for a student

- if they are, they are new or continuing
- which module are they on
- how much have they engaged pre module start with the VLE

And

- if they are a new student then how high a workload have they signed up to
- or if they are a continuing student how positive is their prior success within the OU

Associate lecturers were asked to consider contacting students on the basis of the Early Alerts Indicators and they reported students were generally very happy to be contacted. Students were also asked for their views and they had few reservations of predictions being generated and used as the basis for their tutor to contact them.

The pass rates of the students in tutor groups of the ALs involved in the project showed no consistent differences to the pass rates of those not involved in the project.

| | | |
|-------------------------------------|---------|---|
| Layout of report | | |
| Section 1: Aims and scope | 1.1-1.4 | Background to Early Alerts Indicators and this project |
| Section 2: Activities | 2.1 | Overall plan |
| | 2.2 | Recruitment, background and training for Associate Lecturers in project |
| | 2.3 | Contact with students by ALs |
| | | Contact via web site directly with students |
| | 2.4 | Revision to prediction models as a result of AL feedback |
| | 2.5 | Retention |
| | 2.6 | Changes to original plans |
| Section 3: Findings | 3.1 | AL views |
| | 3.2 | Student views |
| | 3.3 | Simplified model for predicted probabilities |
| | 3.4 | Retention |
| Section 4: Impact and dissemination | 4.1 | To ALs and other academic staff |
| | 4.2 | Students |

1 Aims and scope

1.1 The Open University (UK) has developed two systems of predictive modelling to identify students at risk (Calvert, Rienteas, Zendral). The main function of one system was to specifically predict, pre-start on the module, whether a student would pass their module. The system was subsequently extended to also predict if a student was at risk of dropping out of the module before any of the three key fee liability points. The second system was primarily designed to generate weekly predictions on whether a student was likely to hand in their next Tutor Marked Assignment (TMA). The Early Alerts Indicators combine the two systems within one graphical interface and this interface is made available to Associate lecturers and module teams across all undergraduate modules.

1.2 The aim of this eSTEEeM project was to generate advice, from Associate Lecturers, for the wider Associate Lecturer community about how to make the best use of the information provided in the Early Alerts Indicator tool. As a part of this advice the views of students, on the use of the early alerts data by their tutors, were also to be sought.

1.3 The project aimed to recruit a range of ALs from across two level 1 modules and provide them with access to the complete set of Early Alerts indicator data for their students. Both quantitative and qualitative information from the ALs, on their use of the data was to form an integral part of the project.

1.4 The project started in January 2019 and the intention was to produce results to inform the AL community on the 19 J presentations.

2. Activities

2.1 Overall plan

The intention was to supplement the formal approach of the EAI team with the views of a range of Associate Lecturers. The Associate Lecturers would be asked to offer advice to their peers on the use of EAI data and it was hoped that this would carry a practical credibility with other

ALs. The timing meant that the project needed to start in January 2019 and needed to deliver in time for the advice to be acted upon for the 2019 J presentation. This led to a design that:

- involved volunteer tutors on both MU123 and M140 and on both J and B presentations.
- delivered bespoke training on the use of the EAI for the volunteer ALs and creation of a forum space for the ALs involved in the pilot to collectively discuss their views.
- reminded the ALs two weeks before a TMA was due to consider the EAI data for all their students and decide whether to contact any students on the basis of the data.
- collected from ALs, one week after the TMA was due, a record of who they contacted; the contact method; the students' reaction to being contacted and any additional comments they wished to make.
- collected information on the use of EAI data up to and including TMA02 from the B presentation ALs and information on the use of EAI data, post TMA02, from the J presentation tutors.
- involved the appointment of a lead AL to manage contact with the ALs along with collection, collation and summary of their views.
- involved collecting student views from students on the M140 18J presentation.

2.2 Recruitment and involvement of Associate lecturers

2.2.1 The University regulations involve a lighter set of conditions if twenty or fewer ALs are involved in a project and hence SRPP/ HREC approval was not required. The intention was therefore to recruit up to a total of 20 ALs who tutored across J and B presentations for MU123 and M140. Funding from eSTEEem enabled each volunteer AL to be paid one DL day. This was to cover the initial meeting, a training session time for extra administration returns involved in the pilot and time for a final wrap up meeting. One AL was recruited as the lead AL whose role included prompting the ALs when to contact students, collecting and collating the AL feedback and producing a final report agreed with the other AL's. The lead AL was allocated a further 6 days of DL time. Additionally, one DL day was used to "employ" a member of the EAI team to develop bespoke training for the volunteer AL's and to act as a primary information for the AL's about the EAI package.

2.2.2 ALs were asked in December 2018, via the AL forums, if they would be interested in volunteering to be involved in the project. Many tutors actually tutor on more than one of the MU123/ M140 J/B four combinations in the project and this enabled the volunteer ALs to be allocated in such a way that groups were reasonably balanced. Table 1 shows the allocations. MU123 is a much larger module than M140 and in both modules the J presentation is larger than the B presentation.

Table 1: Numbers of ALs and their associated students involved in the project

| | Number of Associate Lecturers in project | Number of students at 25% fee liability point | Number of students in the pilot programme ALs' tutor groups |
|-----------|--|---|---|
| M140 18J | 4 | 910 | 79 |
| M140 19B | 4 | 560 | 78 |
| MU123 18J | 7 | 1961 | 138 |
| MU123 19B | 4* | 1391 | 126 |

* Note two MU123 B tutors had double groups

2.2.3 The ALs involved in the programme were, of course, atypical in that they volunteered, indicating they had both time and inclination to be involved in the project. Overall, however, there were no concerns that the tutors in the pilot were otherwise atypical of the wider population of M140 / MU123 ALs. There were marginally more younger ALs involved in the pilot; marginally more female ALs; and marginally more had less than 10 years of experience. Further details of the characteristics of the ALs involved and of their tutor groups of students are given in Annex 3.

2.2.4 The ALs who volunteered were invited to a project start up meeting in an Adobe Connect room. The purpose of the project was explained in further detail and a tailored training session on the EAI data provided by a member of the EAI team. ALs were also invited to comment / revise a draft version of the spreadsheet that they would be asked to use in the project.

2.2.5. Twice during the project, the ALs were asked to collect a spreadsheet just for their tutor group and just for the upcoming TMA. These spreadsheets were placed on the project forum space and the lead AL reminded ALs to both collect and then returned the completed spreadsheets. **The spreadsheet provided a checklist of those students whose EAI data the AL was being asked to consider and was also the return summary of the actions the AL had taken.** Figure 1 is an example of the final version of the spreadsheet used in the project and ALs were asked to fill in the yellow columns as appropriate.

Figure 1: Example of the first spreadsheet used by a B presentation AL in the pilot

| Student PI | Student Name (sorted alphabetically) | staff id | Additional contact decided on prior to TMA02 following looking at Early Alert Indicators PLEASE just put a 1 against the name of any you contacted you decided to contact | Successful contact method Phone=1 Email =2 Text=3 | Student reaction to being contacted Generally positive=1 Generally negative=-1 Contacted but could not say=0 Not able to contact=-99 | Additional comments |
|------------|--------------------------------------|----------|--|--|--|---------------------|
| number 1 | name 1 | staffid | | | | |
| number 3 | name 3 | staffid | | | | |
| number xx | name xx | staffid | | | | |

2.3 Contact with students

2.3.1. During the pilot period, which was February to May, each AL had two TMAs due from students. For the B presentation ALs these were TMA01 and TM02 and for the J presentation AL's TMA03 and TMA04. Associate Lecturers often decide to contact their students prior to a TMA and this project was interested in whether, after considering the EAI data, an AL might decide to contact **additional** students. Any such additional contacts were recorded on the spreadsheet in figure 1 and this was then returned to the forum. The lead AL collected the spreadsheets from the forum and collated and shared findings.

2.3.2. The ALs were asked to record their perceptions of how the students reacted to the contact. This was important because it was felt student reaction was a factor in advice to other ALs. The corresponding story, of what the actual students contacted themselves felt, was too complex to gather within the original project. It was, however, possible to gather how students felt in "theory" about using the EAI data as a basis for contact from their tutor.

2.3.3. A voluntary, anonymous questionnaire consisting of just three questions was designed and placed within the M140 J module web site. Students on M140 were already aware that the University held data on their predicted likelihood of success as they had already completed a

TMA01 question that uses the predicted probabilities of passing M140. Additionally, as the students were on a statistics module, it was felt very appropriate to place a voluntary, anonymous questionnaire within the M140 J website. The questionnaire was accessed via a link from the following message:

“I am Carol and I chair M140 J. I would really appreciate you offering me some feedback about an idea we are considering. Any answers to this questionnaire are totally anonymous and will be aggregated and placed back on this site which is good statistical practice.

TMA01 Q5 explained that the Open University has developed some statistical models that can generate information on how likely registered students are to complete all the assignments for a given module. The question showed a boxplot of all students on M140 and you were asked to indicate where you would place yourself on the boxplot.

The information from those statistical models, which the boxplot was based on, is available to your tutor. **How would you feel about your tutor using the statistical model information to help them decide whether to contact you about your studies?** So that might mean your tutor can see that the statistical models are showing that you may be at risk of not completing so they might call you to see if everything is okay. “

2.4 Revision to prediction models as a result of AL feedback

- 2.4.1. As a result of the AL comments a further round of modelling was undertaken. This was undertaken on the part of the system that had originally been designed for predicting success at a point before the course started. The original basis for this model (Calvert) was a logistic regression based on data across all undergraduate courses. In the original model logistic regression was used to identify the best set of factors that predicted passing the course. This set of factors were then applied to the data for students the next year and hence a probability generated that each student would pass their course. The factors and their relative weights were recalculated each year. The predictive value of success for a student was thus based on the experience of “similar” students in previous years and the factors identified to define “similar” were the same across all courses. This model will be called the initial model.
- 2.4.2. The revised model involved only using the data for the previous year on the **same** course. This meant that the variables and their weights could vary from course to course as well as from year to year. This was to address a specific concern of ALs that the key factors effecting success might be quite different on say Mathematics to those for students studying say French.
- 2.4.3. Both the revised and the initial model were used to estimate the impact on accuracy of trying to estimate success at different time points prior to course start. The accuracy figures presented in the findings section of this report are calculated based on pre module start data because ALs were emphasising the importance of information at module start i.e. accuracy is

number of students correctly predicted as fails + the number of students passes correctly predicted as passes

number of students starting the module.

2.5 Retention

2.5.1 This project was not about retention. However, given the importance to the University of retention, it was considered whether there was any impact of the pilot on retention. The pass rates of the students in the tutor groups of the volunteer ALs were compared to those whose tutors were not involved in the project. The predicted probabilities of passing for the students gave a measure of how students were expected to perform – effectively a measure of prior ability. This enabled variations in prior abilities of different tutor groups to be allowed for in the comparison of pass rates. As the B presentation results are not available until September 2019 it was necessary to reschedule the final report and use the interim report to inform the 2019 October presentation.

2.6 Changes to original plans

2.6.1 The original plan was to look at AL usage of the VLE data on the M140 B presentation. This plan was extended when the Early Alerts data, containing the VLE data, was released in October 2018 via tutor home to Associate lecturers. This meant that the VLE data was made routinely and readily available after the project was initially approved. Communication of the VLE data, and its context, to the ALs became much simpler and hence it became possible to extend the project beyond one presentation of one module. The project also extended from the VLE data to all the information in the Early Alerts indicator package.

2.6.2 The clarity with which the ALs consistently emphasised the importance of early information was considerable. There is a trade-off between timeliness and accuracy of the predicted probabilities of success contained in the EAI data and the relationship between the two needed to be explored. The ALs also emphasised the importance of simplicity and transparency of the variables used to generate early predicted probabilities of success. The project was therefore extended to compare the predicted probabilities at module start using the standard approach from the Strategy office and a modified, simplified approach.

3. Findings

3.1 AL views

3.1.1 The ALs felt strongly that the EAI data should be used in combination with, not as a replacement for, other sources of information. Feedback very strongly made the point that there is no substitute for a tutor's knowledge of a student's personal circumstances, ability and commitment. However, it is reassuring for that knowledge to be confirmed by the predictions of student success – students identified as 'at risk' were often those whom the tutor has already decided to contact. While tutors found the data quite easy to use, some initial training was required, possibly together with a further training session once tutors have some experience of using the system. The data is probably more helpful near the start of a presentation, when the tutor is less familiar with students, than later on. While all information is good, there is a danger of overload – use of the data should certainly not be made obligatory, and a tutor's judgment can override the predictions from the model.

3.1.2. AL contact with students may be summarised as:

- Tutors chose to contact between zero and 5 students per group following their scrutiny of the data; however, many added that they would have contacted some of those students even without access to the EAI data.
- The numbers of students contacted decreased with progression through the module, with many tutors choosing not to make any contacts in advance of TMA04.
- Method of contact was more or less evenly divided between email and telephone; a common model was to begin with an email, then follow up with a phone call if the student did not respond.
- Without exception, students who responded to the contact did so positively, indicating that they welcomed a conversation with the tutor. This is a useful finding; some ALs might feel wary of what could be construed as 'hassling' a student, but the study suggests that few if any students will see a contact in that light.

Further details in Annex 1

3.2 Student views: In module questionnaire:

3.2.1. Students on M140 were asked for their views on the use, by Associate Lecturers, of the EAI data and specifically on the data on the likelihood the student will pass the module. Students provided free text on their views, but they were also asked to summarise their views into one of three categories. At the time of that the survey was made available there were 910 students on the M140 J presentation and 69 students responded. Annex 2 contains the free text comments from the students and in figure 2 below the student views are summarised.

Figure 2: Summary of student views on use of predictive data by their tutor to contact the student

Students can always say to their tutor that they prefer not to be contacted. Sometimes "life" happens and throws you an issue and that impacts on your study. If your tutor knows they can sometimes offer ideas of what the OU can do to help you manage your study but often students hesitate to contact their tutor. If the tutor happens to contact the student the issue gets discussed and maybe solutions found.

Please can you choose just one answer from those below that broadly sums up your views.

| Response | Average | Total |
|--|---|-------|
| 1. I am broadly in favour of my tutor using the information to help my tutor decide whether to contact me. |  59% | 41 |
| 2. I do not mind if my tutor uses the information to help my tutor whether to contact me. |  39% | 27 |
| 3. I am broadly not in favour of my tutor using the information to help my tutor decide whether to contact me. |  1% | 1 |
| Total responses to question |  100% | 69/69 |

3.2.2. Both students and AL's reported favourable attitudes to proactive contact from ALs to their students. The basis for the contact was relatively immaterial with ALs more confident about proactive contact if the Early Alerts data was effectively re enforcing the views they had already formed.

3.3 Simplified model for predicted probabilities

3.3.1. The revised model treated new and continuing students separately and it considered MU123 as totally separate from M140. Similarly, the J presentations were also modelled as totally separately from the B presentations. Table 2 summarises the accuracy of both the initial and revised models. The pattern is the same in both cases; greater accuracy is achieved by using the revised model and by treating new and continuing students separately. The results suggest that we should be identifying separate variables and weights for **each course** rather than using a set of variables common across the undergraduate set of all course.

Table 2: Accuracy of predictions prior to module start : Two first year courses using 2016 J data was used to predict 2017 J students results

| | MU123 new students N=891 | MU123 continuing students N=1356 | M140 new students N=315 | M140 continuing Students N=727 |
|---------------|---------------------------------|---|--------------------------------|---------------------------------------|
| Initial model | 61.9% | 72.3% | 69.8% | 70.5% |
| Revised model | 72.3% | 73.3% | 72.1% | 75.5% |

3.3.2. However, the variables identified for new students on MU123 and M140 are the same but with different weights and hence contributions to the logistic regression. This implies that our best predictor variables, for use prior to module start, for new students are simply

- how much the student is studying i.e. total number of credits on registered modules (at PA status at module start) in the current academic year

and

-how much the student engages pre module start with the VLE i.e. number of student module VLE visits per day from prestart to Presentation start

3.3.3 For continuing students the position is, at first sight different, with eight variables that are significant in predicting whether a student passes the module for M140 and ten for MU123. However, just using the four common factors gives over 70% accuracy and simplifies the picture hugely:

- best overall module score from previous study

- total number of previously passed modules (up to 10 years prior)

- total number of previously failed or withdrawn modules (up to 10 years)

--how much they engage pre module start with the VLE i.e. number of student module VLE visits per day from prestart to Presentation start

3.3.4 The process for generating the probabilities of success for a student, pre module start, can then be summarised as considering if the student

- is new or continuing
- which module are they on
- how much have they engaged pre module start with the VLE

And if they are new then how high a workload have they signed up to

or if they are continuing student how positive is their prior success within the OU

3.4 Retention

3.4.1. Whether a student passed a module is highly correlated to the probability that was calculated for the student of passing at module start. Thus, aggregating the predicting probabilities of success of the students of ALs in the project, and comparing it to that for the tutor groups of ALs not in the project, gives a measure of whether the results of the two groups should be similar. Table 5 in Annex 3 shows that there is very little difference between the groups with, if anything, the tutor groups of the ALs in the project tending to be marginally weaker. So the pass rates, based on prior ability, of the students are expected to be similar regardless of whether the students' AL was in the project or not.

3.4.2. For the J presentation the contact with students was prior to TMA03 with many tutors deciding the EAI data gave no reasons to contact students additionally prior to TMA04. If there was any impact on retention of any additional contact by tutors it would be expected on the B presentation rather than on the J presentation. Table 3 actually shows that there is considerable variation in pass rates but no systematic effect within the B presentation.

Table 3: Pass rates by module, presentation and whether in the project or not

| | AL in project | AL not in project |
|---------|----------------------|--------------------------|
| M140 J | 69.0% | 70.9% |
| MU123 J | 61.2% | 67.8% |
| | | |
| M140 B | 61.8% | 53.3% |
| MU140 B | 64.0% | 66.0% |

Note these rates are based on the numbers at module start rather than the more usual rates based on numbers at 25% fee point

4. Impact and dissemination

4.1 To ALs and other academic staff

4.1.1 The impact of this project on student learning and on AL practice cannot be assessed within this project. This project has established a clear message for other ALs – the Early Alert Indicators are useful, but they are a supplement not a replacement to your knowledge. With a secondary clear message that students like you to proactively contact them! The work on simplification should help make the information more acceptable to tutors and the work on quantifying how much accuracy is lost by providing early predictions should help an AL decide how much reliance to place upon it.

4.1.2 Using the interim AL report in the Annex this message has been shared with module teams and Level 1 chairs and thereby influenced the information in use in the October 2019 module presentations. Information about the project has been presented and discussed at an AI development weekend on 7th/8th March 2020. Additionally, an abstract has been supplied for the HEA horizons conference in July 2020.

4.2 To students

4.2.1. There is potentially a clear message for students as well. To a new student it is simply that being cautious over how much you commit to at first is good and engaging with the VLE as soon as you can is good. And that advice is based on empirical evidence from last year's students. To a continuing student the message is similar: - engaging with the VLE as soon as you can is good and bear in mind your previous track record because it is likely to be similar unless you change something.

4.2.2 In both cases we can measure VLE engagement pre module start but actually any engagement pre module start is likely to be positive. For Mathematics and Statistics modules, where we are increasingly running early access to materials.

4.2.3 The most effective way to disseminate that message to students is currently unclear

List of deliverables

None currently

Figures and tables

Table 1: Numbers of ALs and their associated students involved in the project

Table 2: Accuracy of predictions prior to module start: Two first year courses using 2016 J data was used to predict 2017 J students results

Table 3: Pass rates by module, presentation and whether in the project or not

Table 4: Breakdown of AL characteristics

Table 5: Prior abilities of tutor groups in project

Figure 1: Example of the first spreadsheet used by a B presentation AL in the pilot

Figure 2: Summary of student views on use of predictive data by their tutor to contact the student

References

Calvert, C., (2014). Developing a model and applications for probabilities of student success: a case study of predictive analytics. *Open Learning: the Journal of Open, distance and e learning*, 29(2), pp.160-173.

Herodotou, Christothea; Hlosta, Martin; Boroowa, Avinash; Rienties, Bart; Zdrahal, Zdenek; Mangafa, Chrysoula. *British Journal of Educational Technology*. Nov2019, Vol. 50 Issue 6, p3064-3079. DOI: 10.1111/bjet.12853.

University approval processes

Not required as the number of ALs was 20 or under

And

Students views only requested via the module web site and specific to the module therefore SRPP not required

Appendices

Attached at back of report.

Annex 1

Title: AL report on use of early Alert Indicators data

1. Introduction

The Open University holds a great deal of data about its registered students; in some ways the university is better placed than 'face-to-face' institutions in that much of a student's engagement with the university takes place online and is therefore easier to log and to track. The university became aware of this somewhat underexploited resource of data a number of years ago, and the OU Analyse tool was therefore developed. This is a model which combines demographic data about a student with information about the student's use of the VLE, to predict (a) for the next TMA, the risk of non-submission and the likely mark, and (b) for the entire module, the likelihood of completion and of passing. The data is updated weekly, and the model takes account of module-specific factors such as TMA submission dates. The goal of the development of the tool was stated as '[to] identify students at risk of failing the module as early as possible so that OU intervention is efficient and meaningful'. [1]

The tool has been in use for some years on selected modules, but the university now wishes to exploit it more fully to address the issue of retention; it has therefore been subsumed within a broader project, the Early Alerts Indicator project, of which the present work forms a part. The objective at this point is to promote use of the tool to enable effective intervention in cases where students are predicted as at risk of failure.

2. The present project – purpose and methodology

The specific aim of the present project is to capture the experience of a sample of ALs in using the analysis tool, and hence to generate advice for the Associate Lecturer community as to how to make best use of the information provided by the tool.

The project involved a sample of ALs teaching on the 18J and 19B presentations of modules MU123 and M140. All ALs teaching on these modules were invited to take part, and a total of 19 participants was recruited. Volunteers attended an online training session in the use of OU Analyse and were given access to the data for their current group of students on the relevant module.

They were then prompted, two weeks before a TMA cut-off, to visit the OU Analyse site and, on the basis of the current data, to decide which, if any, students they would identify as benefitting from a contact, either by e-mail, telephone or text. They were asked to record:

- (a) Which students they contacted
- (b) The contact method
- (c) The student's reaction to being contacted (positive, negative or unable to contact)
- (d) Any additional comments they wished to make.

It was clearly understood that the activity prompted by the data was in addition to the actions they would normally take, based on, for example, their knowledge of a student's personal circumstances, health issues etc.

Timing of the project was such that TMAs 01-04 were covered for both MU123 and M140 – two TMAs from each presentation. A forum was set up in which participants could share experiences and raise queries. A further online meeting was held towards the end of the project, to gather participants' experiences and views of their use of the data. As a separate exercise, students were also invited to give their views on how they would wish the data about them as individuals to be used.

A number of anomalies with the data, and caveats about its use, came to light in the course of the project; these will be discussed below.

3. Findings

- Tutors chose to contact between zero and 5 students per group following their scrutiny of the data; however, many added that they would have contacted some of those students even without access to OU Analyse, as a consequence of other information available to them, such as the date of last log-in, failure to submit an iCMA, etc.
- The numbers of students contacted decreased with progression through the module, with many tutors choosing not to make any contacts in advance of TMA04. In some cases this was because no students were identified as 'at risk' (those who were having presumably withdrawn or deferred before that point); in other cases, tutors were aware of students' deliberate intention to omit TMA04 and focus on the EMA (neither of the modules in the study has an examination).
- Method of contact was more or less evenly divided between email and telephone; a common model was to begin with an email, then follow up with a phone call if the student did not respond.
- Without exception, students who responded to the contact did so positively, indicating that they welcomed a conversation with the tutor. This is a useful finding; some ALs might feel wary of what could be construed as 'hassling' a student, but the study suggests that few if any students will see a contact in that light. (See also the Appendix for more details of students' views on the use of the data.)
- A high proportion of students contacted welcomed the offer of an extension, despite the fact that they had not requested one. This raises a question as to whether students are given adequate information about extensions – some, for example, appear to believe (notwithstanding the information given via module websites, handbooks etc) that having an extension may result in getting a lower mark.
- Other actions taken as a result of contact with students included arranging one-to-one tutorials and giving advice on deferral. A number of students who failed to respond to contact were referred to the SST in the usual way.

4. Caveats concerning the data

As mentioned above, a number of limitations of the data were noted as the project progressed. These included the following:

- VLE 'clicks' are used as a proxy for student engagement. This is not always reliable, at least for students on mathematics modules, where paper submission is still permitted. A student can log in to download and print a TMA, then do all their study using the print materials, and submit the TMA on paper. This, apart from completing iCMAs online, could be all the VLE activity required for a student to obtain a good result, but students who study in this way may well be identified as 'at risk' due to lack of activity.

- Certain categories of student do not appear in the data- these include students in secure institutions and some late registrations.
- Students who have withdrawn, either formally or in effect, still appear in the data and are likely to show as 'at risk'.
- The algorithm does not 'know about' extensions, and so at certain times a student may appear as 'not submitted' when in fact they have an extension.
- Some 'at risk' identifications which appeared to tutors to be anomalous seemed to be based on the fact that students are taking more than one module. Are such students at greater risk, since they have heavier demands on their time? How, if at all, does their performance on other modules impact on predictions for the module under consideration? This was not clear.
- The system sometimes appears to fail to record when a student has submitted an iCMA, with consequent unjustified 'at risk' predictions. This is under investigation by the OU Analyse team.

5. Tutors' experience in using the data

The ALs involved in the project all found the data both interesting and useful. They provided the following observations and advice about its use:

- It should be used in combination with, not as a replacement for, other sources of information, for example disability disclosure, tutorial attendance, whether or not a student has collected the last TMA, etc. (none of which are incorporated into the algorithm). This was undoubtedly the strongest message from all tutors involved.
- The point was also made very strongly that there is no substitute for a tutor's knowledge of a student's personal circumstances, ability and commitment. However, it is reassuring for that knowledge to be confirmed by the predictions of student success – students identified as 'at risk' were often those whom the tutor has already decided to contact.
- The summary information available for the whole group, and the comparison with previous presentations, were felt to be useful as providing a baseline with which to compare individual student data.
- M140 is already using a system where 5 hours of tutor time are devoted to contacting students near the start of the module; this was felt to be more productive than simply waiting for a student to be flagged as 'at risk', as it can pre-empt problems arising.
- While tutors found the data quite easy to use, some initial training was required, possibly together with a further training session once tutors have some experience of using the system.
- The data is probably more helpful near the start of a presentation, when the tutor is less familiar with students, than later on (this ties in with the fact that many tutors did not feel it necessary to contact any students before TMA04). It was pointed out that students sometimes make a decision to omit the last TMA in order to concentrate on getting the EMA completed on time.
- The timing of the interventions, two weeks before the TMA cut-off, was considered to be about right.
- It would be very helpful if information about extensions could be incorporated into the model (although the message from the team is that this can't be done).
- While all information is good, there is a danger of overload – use of the data should certainly not be made obligatory, and a tutor's judgment can override the predictions from the model.
- The instances where the data seemed very out of line with a tutor's knowledge of a student do not help the credibility of the model, even when the tutor understands why there is a discrepancy.

6. Recommendations

- (i) Tutors should be given access to the data for the groups they are teaching, preferably in advance of the start of the module.
- (ii) There should be training before starting to use the data, with a follow-up session after first use.
- (iii) Students identified as 'at risk of not submitting' should be contacted, if the tutor judges this to be appropriate, approximately two weeks before the TMA cut-off date. Probably the best strategy for contact is email, followed by a phone call if the email elicits no response.
- (iv) Actions to be taken when a student responds to contact are at the discretion of the tutor, but it appears that the offer of an extension is often welcomed (perhaps with some explanation that this will have no impact on marks, but a caution against getting too far behind schedule).
- (v) Students who fail to respond can be referred to the SST in the usual way.
- (vi) There will be a need for some kind of support or reference point for tutors, as queries about oddities in the data are bound to arise.
- (vii) It will of course never be possible to ascribe improvements in retention specifically to the implementation of this process; however, for purposes of evaluation it might be useful, once the process has bedded in, to examine retention and completion rates for the modules involved to see if there has been any improvement.

7. Questions arising

- Should it become standard practice (irrespective of use of the use of OU Analyse) for tutors to contact all students at fixed points in a module – say, at the start and before TMAs 02 and 04? (This would reflect the practice, common in face-to-face institutions, of requiring students to see their tutor at least 3 times per year.)
- Is greater clarity about the use of extensions required as a matter of general university policy?

Lead AL: Clare Morris

ALs involved in project; Ruth Neal, Rowan Thomas, Luay Salman, Clare Morris, Cally Brooker, Julie Owens, Tatianna Macfarlane, Melanie McCabe, Delyth Tomas, Val Aspland, Iain Chapman, Jim Thorpe, Sandy Barthorpe, Julie Youle, Barry Goodman, Col Blundell, Dave Brown, Felicity Bryers, Cheryl McAndrew

Annex 2: Students free text responses to use of Early Alert Indicator Data by ALS

1 How would you feel about your tutor using that information to help them decide whether to contact you about your studies?. (the free text box below is 10000 characters so you can say quite a lot)

Response

The tutor role is to offer support and guidance and any tool that helps them to this can only be a good job, also some individuals don't always find it easy asking for help, especially if they feel they are asking more than offers, so their tutor contacting them and offering advice might be a welcome support

i would have no problem with my tutor contacting me about my studies

Fine

Excellent idea. I would have no problem with contact. Support is support.

Fine

I wouldn't have a problem with that, it sounds like an effective way to target resources at people who potentially need more help.

No problem - sounds like a helpful idea

Personally I would have no problem with my tutor contacting me about my studies. I might be concerned that the contact would indicate a problem rather than an opportunity to have some timely guidance.

I'd be more than happy to allow this. Due to the nature of distance learning and, in my own case fitting it around a full time job, sometimes support is necessary. I see no problem in being contacted by the OU if they feel I need it.

I think that it is a good idea as it allows the tutor to check up on certain students who need the help, and can let the tutor know if there was a certain issue impacting their studies. This can help the student get back on track easier.

I'd be ok with a tutor contacting me.

Yes I think it is a good idea. They are here to help us and if they see that we might need help they could use it to nudge us in the right direction.

No problem at all.

Fine

I would feel fine about this, since the aim is not to intrude on the student's privacy but to offer them help that may be available that they may not know about, and since the aim is to assist the student in achieving their best from the module.

I would be happy about that, as long as my tutor took no for answer if I refused help at a particular time. I have once or twice got a long way behind (in other modules) for life reasons but caught up quickly at the appropriate time. Multiple attempts to contact me would not have helped...

That would be good as additional support could be offered that we are unaware of

No strong feelings either way

I would very much like that my tutor uses the information provided to him and contacts me immediately if he sees that I've not been doing my work. This will help others students as well because they might lose track of their assignments and derail to a down-slope.

I would be happy that they contacted me .

I'm quite happy for my tutor to use statistical information to contact me. I'm assuming that the data would be only one part of any evaluation that a tutor uses to decide whether or not to contact a student. It would be more important to me to have an ongoing dialogue with my tutor through out my module so that any issues, that may present a risk to my successful completion, can be flagged early and dealt with.

I would not mind being contacted by e-mail if there were concerns. This would give me an opportunity to open up to a tutor about any issues I would be having at the time. As it is in e-mail format, it would also be easy to politely decline any offer of support / help if I did not feel I wanted any at that stage. I would personally prefer to be contacted via e-mail rather than a phone call where I might feel I was put on the spot.

I think this is a good idea because some learners will not directly contact the tutor for fear of embarrassment/lack of knowledge/time etc so for a tutor to reach out to learners in the lower quartile, I would think that is acceptable.

I wouldn't mind my tutor contacting me anytime about anything concerning my studies.

I think that would be fine.

I think that would be fine.

I think it would be a good idea. With limited (or even no) face-to-face contact, it must be difficult for tutors to identify if individual students may need help. They are reliant on students contacting them but what if the student doesn't feel confident enough to initiate the contact. The more information a tutor has available, the better they will be able to judge how their students are performing.

I have no issues with this

Personally, I'd be happy, but essentially my reply to the question was "I'm pretty confident I'll be fine". In terms of a conversation starter it works, though maybe my concern would be that it might discourage people who were nervous about the course.

I think so long as people know it is the tutors trying to help and check you are ok rather than being targeted for not performing well, it is fine.

I would be OK with the use of the model providing that the tutor used the information sensitively and the contact is clearly focused around helping me with any issues I might have.

I feel contacting students when they are lacking confidence or struggling is a good idea. I'm sure many students need help they just do not want to ask.

I would feel absolutely fine if my tutor contacted me about my studies, sometimes you need a little encouragement and an understanding ear. I can only see positive outcomes from having someone looking out for you.

I think it's a good idea for tutors to check in regularly with all students, but if they knew info like this, they could make sure that students that have the potential to struggle more are reminded that help is available. I have one tutor that checks around TMA time and another that doesn't. I struggle to contact both just to check in when I'm doing good. Having tutors make the move to contact regularly is welcome.

I don't mind.

I would like to, being contacted by my tutor about possible low on the performance of my studies could help me to improve.

I wouldn't mind if the tutor used that information.

I don't have any objections to the tutors using the information. I am not sure how useful it is going to be, as I would imagine people would presumably score themselves higher to not look like they are struggling, especially if they thought the tutor could see their answers and contact them because of it.

I am happy for them to have that information and use it as they see fit. I tend not to contact my tutors directly, but as long as the tutor wasn't pushy I wouldn't mind being contacted.

If it is to help me complete my studies and offer support at the appropriate time then I would welcome the idea.

That's fine but I would expect to be told that it might be used for that before I answered it.

I'd be ok with that. I haven't needed to ask my tutor for help yet as any questions I've had, I've been able to ask at the online tutorials. I do feel that if I did need to speak to my tutor, I would feel awkward to do so.

Absolutely fine.

That is ok with me. Anything that could help your tutor help you.

I think it is a good idea. It's a chance for people to say they don't feel comfortable and could do with the extra support or vice versa.

I think it is a good idea but feel that student that think they have a command of the subject may be left out of tutor's attention

Yes I think that would be a good idea.

I don't see any reason for a tutor not to contact a student. Use of that sort of information would be reasonable. Should be via email initially.

Far better to treat all students equally, and not to try to tailor services on a rather arbitrary analysis. Those of us who think we might struggle may also be the ones that need to be left alone to get on with things, knowing the tutor is there but without feeling "pestered".

If my tutor was worried that there was a problem I would prefer to be able to either clarify that there isn't one, or make them aware if there is and how it may affect my studies.

I'm fine with that.

I would feel OK about this, as long as this policy was made clear and opt-out.

I'd be very happy with this

I have no problem with that

I would have no objection.

I think it will be great. It might be a lifesaver and help me when I thought I was going great. If at the end it results that it was all a false alarm no harm done.

It could be helpful to intervene if the student has not raised an issue but it could be offputting if there is no problem despite what the data is saying. So I'm mixed on it.

I'm comfortable with a tutor using this information to recognise if contact may be helpful, but that alone should not be used in determining this. Dependant upon the drivers in the predictive model it may miss those who require further support, but due to past performance sit near or above the mean. The OU allows you to take a wide range of modules where the subject matter can differ wildly. For instance, I'm doing business modules along with a stats module. Those other models are related in subject matter, whilst stats is not. If my historic performance in other related modules drive the predicted performance to my stats module, and therefore not flag as 'at risk' to my tutor then an opportunity to recognise and provide early support may be missed, should it so be needed.

I don't have a problem with that as if I was lagging behind it would be helpful to know and have tutor support

I would be happy for my tutor to contact me in response to that data, or indeed on any occasion concerning my studies.

I would feel more than happy about my tutor using any information available to them in order to offer help when needed.

Worth at least the Tutor asking the question of the student. For many newer students may not be aware of the tutors support role if all they can compare it to is their high school education. A more experienced student may have found it all easy to this point so again asking the question and opening the discussion would be worthwhile. The student can always politely decline.

I feel that would be fine.

Yes, that's fine

I would feel that if help is required, i would like to be contacted

I don't have anything against it. In fact I think it can be helpful sometimes to having someone checking on you in case something seems out of the ordinary.

Not bothered at all. Tutors are there to guide and direct us.

Annex 3 Table 4: Breakdown of AL characteristics

| | Number of ALs (individuals) in project | Number of active ALs on <i>module</i> (individuals) |
|---------------------------------------|--|---|
| M140 | 8 | 59 |
| 59yrs or younger | 6 | 63% |
| 67yrs or older | 1 | 12% |
| Female | 6 | 56% |
| 10 or fewer years since OU start date | 6 | 47% |
| 20 or more years since OU start date | 1 | 17% |
| Average number of modules tutor. | 3.1 | 3.4 |
| MU123 | 11 | 129 |
| 59yrs or younger | 6 | 53% |
| 67yrs or older | 3 | 27% |
| Female | 6 | 49% |
| 10 or fewer years since OU start date | 4 | 29% |
| 20 or more years since OU start date | 4 | 28% |
| Average number of modules tutor. | 3.9 | 3.6 |

The predicted probabilities of passing in table 5 have been calculated using the simplified approach described in section 3.3. Not all the tutor groups will have exactly the same pattern of predicted probabilities, generated at module start, of passing for the students. The average probabilities, shown in Table 4, though indicate that the average predicted levels of success are close between those in the pilot and those not – thus for M140 J the average probability, as generated pre module start, of passing is 71.277 compared to 71.279 for those not in the pilot. If anything, the tutor groups of those AL involved in the project has marginally lower prior probabilities of success.

Table 5: Prior abilities of tutor groups in project

| | Average predicted probabilities of passing for student in the pilot tutor groups | | Average predicted probabilities of passing for student in the non-pilot tutor groups | |
|---------|--|-------|--|---------|
| M140 J | 0.71277 | n=148 | 0.71279 | n=790 |
| M140 B | 0.61364 | n=76 | 0.61935 | n=508 |
| MU123 J | 0.6846 | n=208 | 0.6850 | n=1834 |
| MU123 B | 0.6445 | n=178 | 0.6429 | n=1108. |