

Medical School Performance as A Predictor of Scores in the National Medical Specialization Exam in Turkey

Ahmet Murt (✉ ahmet.murt@istanbul.edu.tr)

Istanbul University Cerrahpasa Faculty of Medicine: Istanbul Universitesi-Cerrahpasa Cerrahpasa Tip Fakultesi <https://orcid.org/0000-0002-1948-2914>

David Hope

University of Edinburgh

Recep Ozturk

Istanbul Medipol University

Helen Cameron

Aston University

Research article

Keywords: Medical school performance, external exams, assessment, validity

Posted Date: March 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-304816/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Medical educators and assessors like to include predictive validity in their validity arguments but relevant evidence may be difficult to find. External standardized examinations may have a role in validating both the educational process of medical schools and their assessment results and outcomes. A strong correlation between medical school and external exam performances may also lend evidence of validity to the external examination. This work from one of Turkey's top medical schools explored the correlations between students' medical school performances and scores from the Specialization in Medicine Exam (TUS). The TUS is a post-graduate national ranking examination.

Methods

A total of 246 students from two different programs of a medical school, which have identical curricula but different admission scores were studied retrospectively. Students' year based Grade Point Averages (GPAs) and end-of-school (graduating) GPAs were calculated using a weighted mean method. Bivariate correlations were calculated between year specific GPAs, graduating GPAs and TUS scores.

Results

Students' inter-year GPAs showed strong significant correlations ("r" ranging from 0.59 to 0.86, $p < 0.001$). Their graduating GPA also had a strong significant correlation with TUS scores ($r = 0.65$, $p < 0.001$). Linear regression models showed the significant relation between medical school performance and post-graduation national exam performance.

Conclusion

Student success has a high degree of consistency throughout the medical school and students' performance across all domains of assessment in the undergraduate program might be a good predictor of cognitive skills in an external national examination in the early postgraduate phase.

Background

Postgraduate external examinations are used in different regions, some for licensing to practice and some for selection purposes. The Specialization in Medicine Exam (TUS) in Turkey is a ranking examination that matches the medical school graduates to available residency programs. TUS is organized twice a year (in April and September) by a central examination council (abbreviated as OSYM in Turkish). There is no limit to the number of attempts a medical doctor can sit TUS, but students prefer to prepare earnestly for their first attempt just following graduation. As graduation is generally in July, the

first attempt to sit TUS is in September. A minority of students who graduate after September, prepare for the April TUS of the following year. The general reason for students to sit the exam in April is a delay for any reason of their graduation from medical school. Some students strategically delay their graduation in order to have more time to study for the exam. In order to be eligible to sit TUS, students must pass the foreign language test. English is the most popular foreign language but candidates may also choose French or German.

TUS comprises two tests (basic medical sciences and clinical sciences), each composed of multiple-choice-questions (MCQs). Basic medical sciences and clinical sciences tests are sat in two different sessions on the same day. TUS as the external assessment of undergraduate studies might provide evidence of the predictive validity of the schools' assessment results and outcomes. Specialty exit exams may also be used as a validation source but as they are too far removed in time from the medical school exams, many other factors may influence their results. As an exam composed of only MCQs with no assessment of observed clinical skills and competencies, the TUS is criticized by authorities that, it may not address the same learning outcomes as the school exams. In contrast, Medical schools, use a wide range of assessment methods that test clinical, personal and cognitive skills in an effort to more accurately assess the performance of students [1]. Thus, if there is a strong association, between performance at medical school and TUS this may also validate the latter as a tool to select the most suitable candidates for residency.

The aim of this study is to seek evidence of the predictive validity [2] of one school's assessment results using both results throughout the medical program and the specified external exam.

Methods

Setting

The study was conducted in Istanbul University-Cerrahpasa, Cerrahpasa Medical School. Undergraduate medical education in Turkey lasts six years. Students learn basic medical sciences in the first three years, attend clerkships and study clinical sciences in years 4 and 5 and have an internship without examinations in the final year. There are 2 medical education programs in Cerrahpasa Medical School, one in Turkish (program 1) and the other in English (program 2). Programs 1 and 2 are identical with the same curriculum and delivered lectures. The only differences between the two programs are the language used, and the fact that program 2 (English) requires a higher admission score. Program 2 of Cerrahpasa Medical School ranks top in Turkey for students' admission scores.

Students and TUS Dates

A total of 330 students, (270 program 1, 60 program 2) matriculating in 2007 were included in the study.

Total completion rate for program 1 was approximately 89% (240/270) across the six-year program. Twenty additional students were accepted into the program from different Turkish medical programs in

different years through the process of lateral transfer. Nine students were unable to pass the foreign language examination and were ineligible to sit TUS. A total of 201 students from the program 1 took the first TUS following their graduation

The vast majority of students from program 1 graduated from medical school before September so their first attempt to sit TUS was in September 2013. There was also a minority of students who graduated from medical school after September and their first attempt to sit TUS was in April 2014. This study included all students who took the first TUS following their graduation independent of their graduation date. Fifty students from program 1 either delayed sitting TUS or decided not to take it.

All final graduating GPAs of the 201 program 1 students who sat TUS were available but one or more interim GPAs were missing for 21 students.

Total completion rate for program 2 was around 92% when all 6 years are taken into account. The lateral transfer process added nine more students to program 2. No students failed the foreign language test. This was expected as all students pass an English proficiency test before admission to the program and all lectures are delivered in English. TUS scores were available for the 45 students who took the first TUS after graduation. Nineteen students from program 2 either delayed sitting TUS or decided not to take it. All final graduating GPAs of the 45 students program 2 who sat TUS were available but one or more interim GPAs were missing for 5 students.

Similar to Program 1, the majority of students from Program 2 graduated from medical school before September and their first attempt to sit TUS was in September 2013. A small number of students whose first attempt to sit April 2014 TUS were also taken into account.

Collection of data and variables

Performance of a graduating student at the medical school under study is assessed using a compilation of a variety of assessment methods and the contributing scores were accumulated into a Grade Point Average (GPA). Weighted means calculated as $[(\sum \text{grades} \times \text{credits}) / \text{credits}]$ on a 4.0 scale gave interim GPAs and graduating GPAs were calculated from interim GPAs. Multiple-choice questions (MCQs) constitute around 70% of the 1st year GPA. This ratio decreases gradually to almost 55% for 5th year. Contribution of interim GPAs to graduating GPA is almost equal for each year.

School grades were obtained from the University student affairs office and scores at the first attempt of TUS (at the end of medical studies in 2013) were provided by the central examination council (OSYM).

Ethics and Statistical Analysis

The study was approved by the institutional review board of Cerrahpasa Medical Faculty. (Approval number: 2014/A-37)

The researchers were blinded to the identity of students and their scores by having the examination council and student affairs teams apply codes to the data before handing them over. This secured the

confidentiality of the data.

Looking at the association between school performance and TUS scores has been proposed as the core of the study. Students from different programs of the school were also compared to check if the correlation is transferrable and was not affected by the fact that program 2 accepts students with higher admission scores.

Statistical analysis was carried out by R open-source package [3]. Descriptive statistics were calculated. Inter-year specific GPA correlations have been calculated in order to check the degree of multi-collinearity if different year GPAs were used together to predict TUS score. This also helped to evaluate the consistency of students' performance across all years of medical studies. A linear regression model was applied to formulate a relation between graduating GPA and TUS scores. Students from both programs were included in the first model looking at the general predictive power of GPA for TUS. As a second step, linear regression models specific to each program were compared.

Results

Medical school GPAs show a normal distribution for the entire medical school period (Graph1). Mean GPA for students for 1st year is 3.01 ± 0.43 , the same as the mean graduating GPA (SD = 0.39). The mean GPA decreases to 2.95 ± 0.52 in 2nd year, which anecdotally from students appears to be the most difficult year of the program. Students' GPA grades show the largest variation in this year being as high as 2.28. The mean GPA for 3rd year is 3.04, which is slightly above the mean of the final GPAs. The mean GPA for 4th year is the lowest of all (2.83 ± 0.45). The mean GPA for the 5th year is the highest of all reaching 3.31. The lowest GPA score from 5th year is 2.45 (SD = 0.35) and the range narrows to 1.55. (Table 1) The mean score for TUS, for all 246 students is 56.8 ± 10 .

GPAs of each year (from 1st to 5th year) showed a medium to large correlation with TUS scores (Table-2). Effect size was the lowest for 1st year and highest for 4th year. Fourth year, as the main year of clinical phase of the medical school was the most decisive year to separate high and low scoring students in the TUS. Graduating GPA had a positive statistically significant correlation with a large effect size of 0.65. The inter-year GPA correlations (Table 2) demonstrate that the pre-clinical years (1st, 2nd and 3rd) correlated better with each other than with clinical years. Similarly the clinical years (4th and 5th years) correlated better with each other than with the first three years.

Year 1 GPAs correlate moderately with TUS Scores ($r = 0.44$) and has the lowest correlation among all the year specific GPA correlations with TUS. This can be explained by the fact that Year 1 includes assessment of foreign language (English), Turkish language and Turkish history, none of which is required for the medical profession nor form part of the TUS examination. Second year GPAs correlate with TUS better than 1st year ($r = 0.51$). Correlations of year based GPAs with TUS show an increasing pattern in 3rd and 4th years.

Graph 2 with two different dashed lines indicates that the pass/fail score of TUS (45) correlates with the GPA of 2.75. A GPA of 2.75 can be claimed to be the cut-off to identify students at risk of failing TUS. No student who has a 4th year GPA higher than 2.75 failed TUS. This potentially useful indicator needs to be further tested looking at different cohorts of students.

All 246 students were included in the linear model that regresses TUS scores on the final GPA. The final GPA made up 44% variance of the TUS scores. The model was found to be significant with a $p < 0.001$. The graphical illustration of the linear model (least squares line) is given in graph-3 and indicates the pass score for TUS (45) with a horizontal dashed line.

We analysed the correlations between medical school performances with TUS scores separately for each of the Program cohorts. For the 201 students admitted to Program 1, (lower university admission scores), the mean final GPA was 3.02 ± 0.40 , while the mean TUS score was 56.61 ± 10.28 . The model was significant ($p < 0.001$) in defining the association between GPA and TUS scores. The graduating GPA accounted for 42% of variance in the TUS scores. The mean graduating GPA of students ($n = 45$) in program 2 was $3.09 \pm SD = 0.41$ and they had a mean TUS score of 58.57 ± 9.60 . According to the linear model, 51% variance of TUS scores could be explained by students' graduating GPAs. The linear model was significant with a $p < 0.001$. Medical school performances, national exam scores and the correlation between both did not show any statistical difference for two different programs.

Discussion

In our study, TUS, a national, external ranking examination-comprising MCQs only, is strongly correlated, with medium to large effect sizes, with students' grades at medical school. Although TUS does not specifically assess the demonstration of clinical skills, the results of TUS are strongly correlated with the school results that incorporate tests of practical competence as well as cognitive skills. This may be due to the underlying factors (intelligence, conscientiousness etc) that determine performance. Although this study demonstrates that a student's performance across all domains of assessment in the undergraduate program is a good predictor of cognitive skills in an external national examination in the early post graduate phase we can not comment on the correlation with early postgraduate practical competence. Had the TUS included tests of practical clinical competence it might not have been surprising to find that after a year-long internship, students' skills had accelerated. Whether that rate of development could be predicted by their total undergraduate performance remains unknown in our context and should be explored in future projects.

Previous similar studies, mainly from North America, use United States Medical Licensing Examination (USMLE) as the criterion variable [4–6]. There are also examples from the Netherlands [7] and Australia [8]. All report similar findings with those from North America. This study helps to generalize the findings about the correlation between undergraduate performances with an external exam internationally. However we also found that students' performance across a medical program is (fairly) consistent in keeping with the findings of Hope and Cameron [9] and McManus et al [10]. McManus et al. identified

that the continuity of academic success for medical students stretched from secondary school into the early years of their postgraduate careers and referred to this as *'the academic backbone'*

Although GPAs from all years have moderate to large correlations with TUS, there is a trend for the correlation power to increase from 1st to 4th years. Year 1 has the lowest correlation of all. This may be partly because some subjects in the 1st year curriculum such as history, foreign language, Turkish language etc. are not directly related to medical sciences and are not reflected in the TUS. The low correlation may also reflect the effect the adaptation period has on some students in an unpredictable fashion. The 4th year medical school GPA is the best predictor of TUS scores ($r = 0.67$). This may reflect the alignment between the medical school curriculum and the focus of the TUS; in this first clinical phase of education year 4 students have the main clinical clerkships such as Internal Medicine, General Surgery, Gynecology/Obstetrics and Pediatrics. The high correlation with year 4 may also reflect the development in students' professional identity and thus their motivation to study. Fourth year may also be the year when all students' study strategies become more focused to plan their careers for specialization. In this study the 4th year GPAs correctly identified students at risk of failing TUS. If this finding is ratified with further cohorts it will provide an evidence-informed metric that the medical school can use to detect and support those students who may need further remediation before they proceed further. The 5th year GPAs are the highest and the grade range narrows to 1,55. This might be due to the checkpoint function of 5th year. As students require a cumulative GPA over 2.00 to pass to the final year; they may be preparing hard for the exams in order to lift their grades.

What message should findings in this study give to policy makers in Turkey and internationally? In the current system, TUS is the only assessment that determines which students enter specialization (residency) training. If TUS is strongly correlated with medical school performance, does this indicate that TUS is a valid and reliable method to select candidates for each residency? There were some students who outperformed their peers in TUS although they had a lower school performance. Such a final, end-of-school exam may be identifying some students who have accelerated their performance during their internship due to improved motivation, a preference for learning in the workplace, unknown factors or a combination of these. On the other hand TUS may be criticized for its focus on MCQs only. In contrast, medical school performance could be considered a better measure of global achieved learning using a range of assessment tools that evaluate problem solving, communication and practical skills [11]. In addition, the medical school's use of multiple methods allows for compensation of each tool's weaknesses [1]. At this stage while the TUS continues as an MCQ only examination it is difficult to determine if the unexplained variance between school and TUS performance is due to the difference in learning outcomes addressed by different assessment tools or by differences in achieved learning over the internship.

It is reassuring that there is a good correlation between year GPAs suggesting that they may each be useful additional determinants for selection to residency slots.

The results suggest that for students in the medical school under study, adding the undergraduate performance to the TUS results may increase the validity of the decisions about the allocation of residency posts. However it cannot be assumed to be so across medical schools. In order to be more confident about graduates' competence and the selection processes, studies looking at further cohorts and across a range of medical schools is required.

A major strength of this study is being able to include all students with available data in the analysis. As the student data were available longitudinally, consistency of student success could also be evaluated. With a strong inter-correlations of year specific GPAs we can claim that student success has a high degree of consistency. This study also has some limitations. It is based in one-institution and analyzes student outcomes from one year of graduation. Further work should cross-validate the results by involving different institutions and graduates from different years. Another factor is that, students make extensive preparations for TUS, which may limit our comments for direct effects of school curriculum on TUS. We couldn't analyze the gender effect on student performance as assigning gender codes to unidentified students could jeopardize the confidentiality of data due to small numbers in the study. Although we have demonstrated that medical school performance is a good predictor of performance in the early graduate phase our study was not designed to explore the causal mechanisms. Do hard-working successful students become hard-working successful graduates? Are the exam results in medical school or in the TUS a result of commercial preparation courses? Does learning in the internship year differ from that in the first 5 years? Why do students fail the TUS? These interesting questions, raised by this analysis, will require a more qualitative approach to address them.

Conclusion

This study demonstrates that medical school performances correlate well across the years of the program and with the Specialization in Medical Exam (TUS) that is a national ranking examination. Student success in medical school shows consistency with a medium to large correlation between yearly GPAs and 42–52% of the variance of TUS scores was attributed to the graduating GPA.

Year 4 GPAs showed the highest correlation with TUS scores thus suggesting a timely metric that the medical school can use to detect and support those students at risk of failing the TUS.

There is no compulsory medical licensing examination in Turkey for new graduates and thus no direct validating assessment. This study explored the national MCQ based examination for selection to specialty and found a strong correlation between medical school performance and TUS for one cohort of students at one medical school. This provides evidence to validate the educational program of the medical school externally but requires further study to generalize the findings to other cohorts and other medical schools in the country.

Abbreviations

GPA
Grade Point Average
MCQ
Multiple-Choice Questions
OSYM
Central examination council
TUS
Specialization in Medicine Exam
USMLE
United States Medical Licensing Examination

Declarations

Funding: Not applicable

Competing Interest: Authors declare no conflicts of interests.

Ethics Approval and consent to participate: The study was approved by the institutional review board of Cerrahpasa Medical Faculty (Approval number: 2014/A-37). Informed Consent was not needed as identities of all subjects were blinded to all researchers. This was also approved by the same review board.

Consent for publication: Not applicable

Author's Contributions: AhM conceptualized the idea, did the literature search, collected the data, performed the analysis and drafted the manuscript. DH designed and performed the analysis. RO was involved in planning and supervision of the work. HC critically analysed the literature, aided in interpreting the results and worked on the manuscript. All authors discussed and commented on the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from central examination council of Turkey (OSYM) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of central examination council of Turkey (OSYM).

Acknowledgement: We would like to express our gratitude to former Cerrahpasa Medical Faculty Dean Prof. Ozgun Enver and former Istanbul University Rector Prof. Yunus Söylet for their unconditional support to the work.

References

1. Epstein RM. Assessment in Medical Education. *New England Journal of Medicine* 2007; 356:387–396
2. Trochim, William M. The Research Methods Knowledge Base, 2nd Edition. Internet page, at URL: <<http://www.socialresearchmethods.net/kb/>> (version current as of October 20, 2006).
3. Ihaka, R. & Gentleman, R. (1996). R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299–314.
4. Elam C.L, Johnson M.M.S. NBME Part 1 versus USMLE Step 1: Predicting Scores Based on Preadmissions and Medical School Performances. *Academic Medicine* 1994; 69(2): 155
5. Paolo AM, Bonaminio GA, Durham D, Stites SW. Comparison and Cross-validation of Simple and Multiple Logistic Regression Models to Predict USMLE Step 1 Performance. *Teaching and Learning in Medicine* 2004; 16(1): 69–73
6. Zahn CM, Saguil A, Artino AR Jr, Dong T, Ming G, Servey JT, Balog E, Goldenberg M, Durning SJ. Correlation of National Board of Medical Examiners scores with United States Medical Licensing Examination Step 1 and Step 2 Scores. *Academic Medicine*.2012; 87: 1348-54
7. Cohen-Schotanus J, Muijtjens AM, Reinders JJ, Agsteribbe J, van Rossum HJ, van der Vleuten CP. The Predictive validity of grade point average scores in a partial lottery medical school admission system. *Medical Educaion* 2006; 40(10): 1012–9
8. Edwards D, Friedman T, Pearce J. Same admission tools different outcomes: a critical perspective on predictive validity in three undergraduate medical schools. *BMC Medical Education* 2013; 13: 173
9. Hope D. & Cameron H. Academic performance remains predictive over a five year medical degree. *Innovations in Education and Teaching International*. 2018; 55(5): 511–520
10. McManus IC, Woolf K, Dacre J, Paice E, Dewberry C. The Academic Backbone: Longitudinal Continuities in Educational Achievement From Secondary School and Medical School to MRCP(UK) and the Specialist Register in UK Medical Students and Doctors. *BMC Med*. 2013 Nov 14;11:242. doi: 10.1186/1741-7015-11-242
11. Veloski JJ, Rabinowitz HK, Robeson MR, Young PR. Patients Don't Present with Five Choices: An Alternative to Multiple Choice Tests in Assessing Physician's Competence. *Academic Medicine* 1999;74(5):539–546

Tables

TABLE-1 : Descriptive statistics of the data

	Mean	SD	Median	Min	Max	Range	Skew	Kurtosis
Year 1	3.01	0.43	2.94	2.16	4.00	1.84	0.35	0.68
Year 2	2.95	0.52	2.88	1.72	4.00	2.28	0.33	0.72
Year 3	3.04	0.48	2.99	1.88	4.00	2.12	0.33	0.69
Year 4	2.83	0.45	2.79	1.75	3.96	2.21	0.25	0.35
Year 5	3.31	0.35	3.35	2.45	4.00	1.55	0.38	0.45
Final GPA	3.01	0.39	2.97	2.14	3.98	1.84	0.38	0.51
TUS(Clinical)	56.89	9.96	58.49	20.86	72.57	51.71	0.82	0.38
TUS(Basic)	56.73	10.0	58.19	14.90	73.32	58.42	0.91	0.94

Table-2: Correlation degrees of year specific GPAs, graduating GPA and TUS scores.

	Year 1	Year 2	Year 3	Year 4	Year 5	GPA all	TUS-Clin	TUS-Bas
Year 1	1.00	0.81	0.74	0.61	0.59	0.81	0.44	0.44
Year 2		1.00	0.86	0.71	0.69	0.89	0.51	0.51
Year 3			1.00	0.78	0.72	0.92	0.58	0.58
Year 4				1.00	0.80	0.92	0.67	0.67
Year 5					1.00	0.87	0.60	0.60
GPA all						1.00	0.65	0.65
TUS Clin							1.00	1.00
TUS Bas							1.00	1.00

Figures

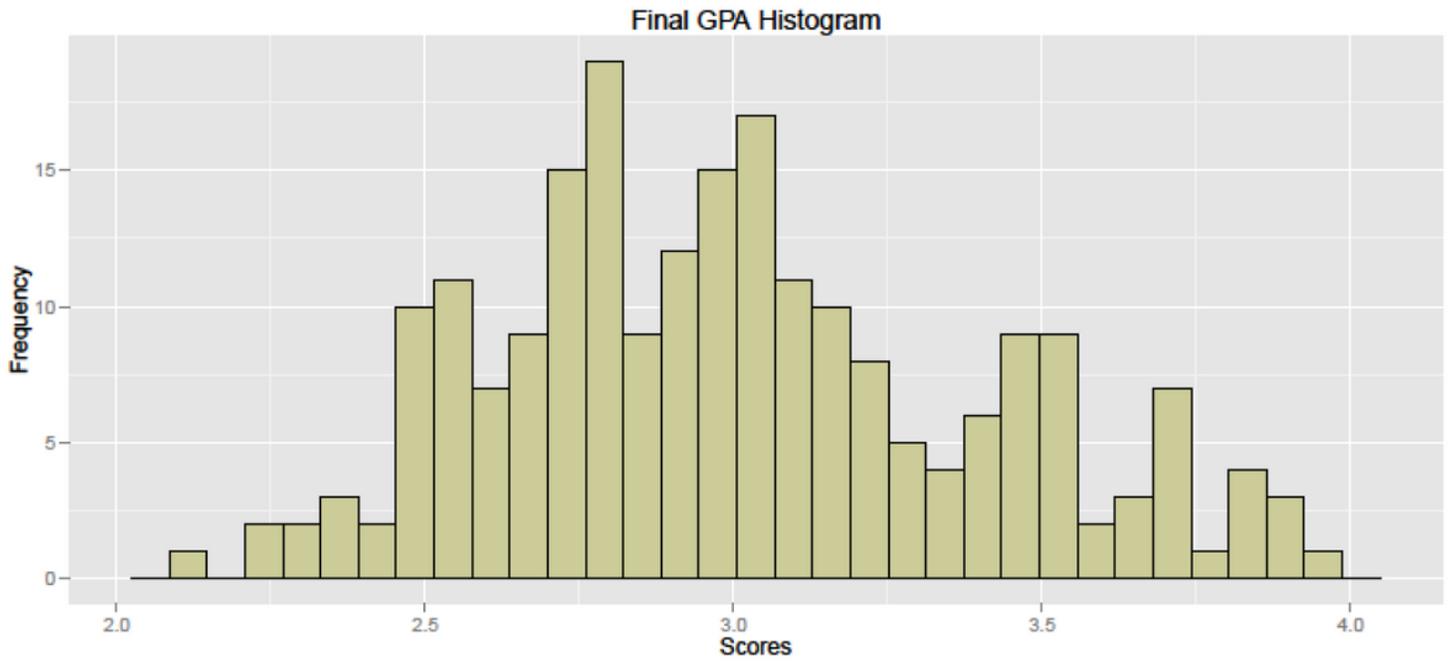


Figure 1

Histogram of final GPAs of all medical students

Year 4

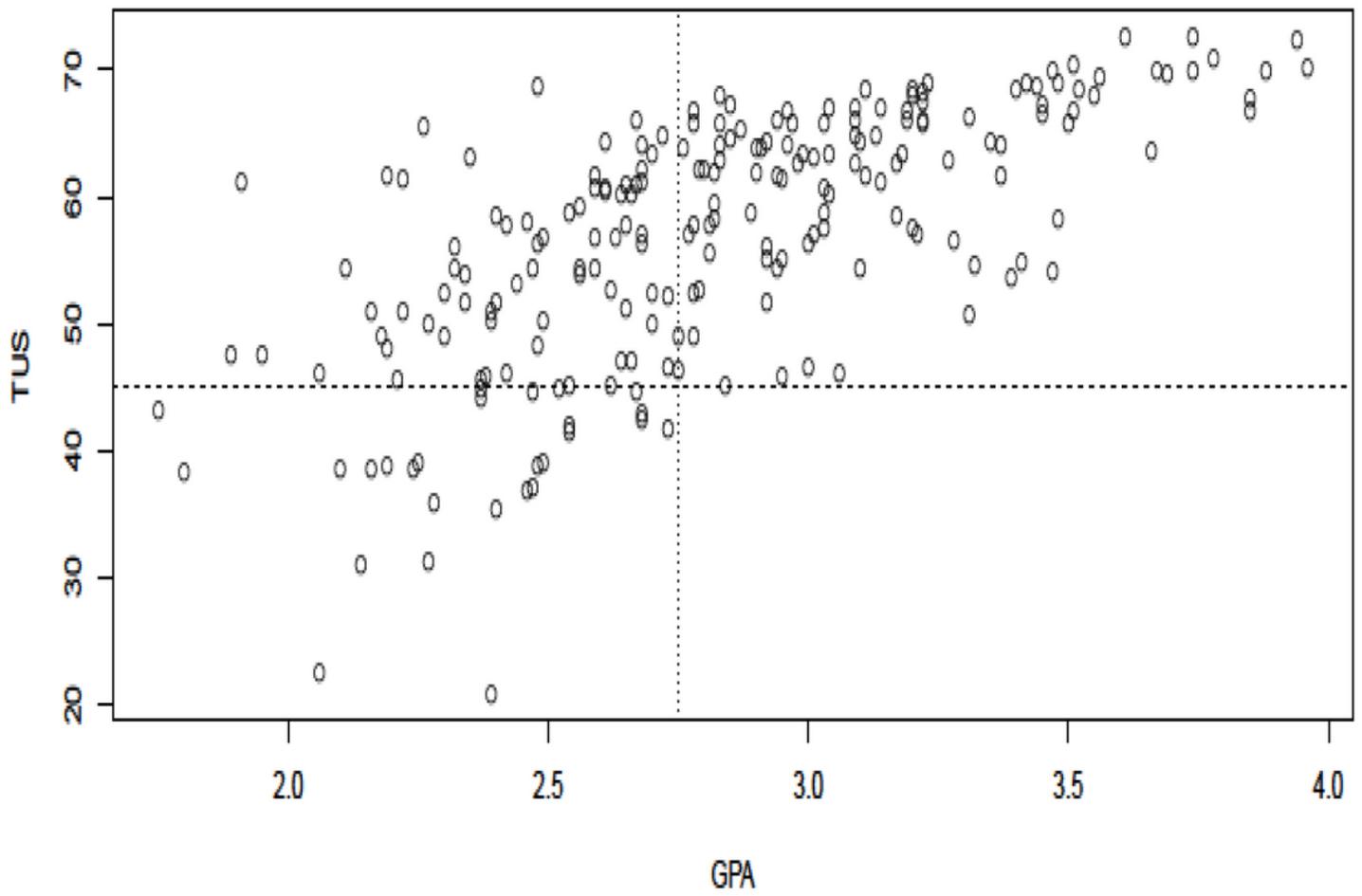


Figure 2

Year 4 GPAs and TUS scores plot by assigning a cut-off GPA of 2.75.

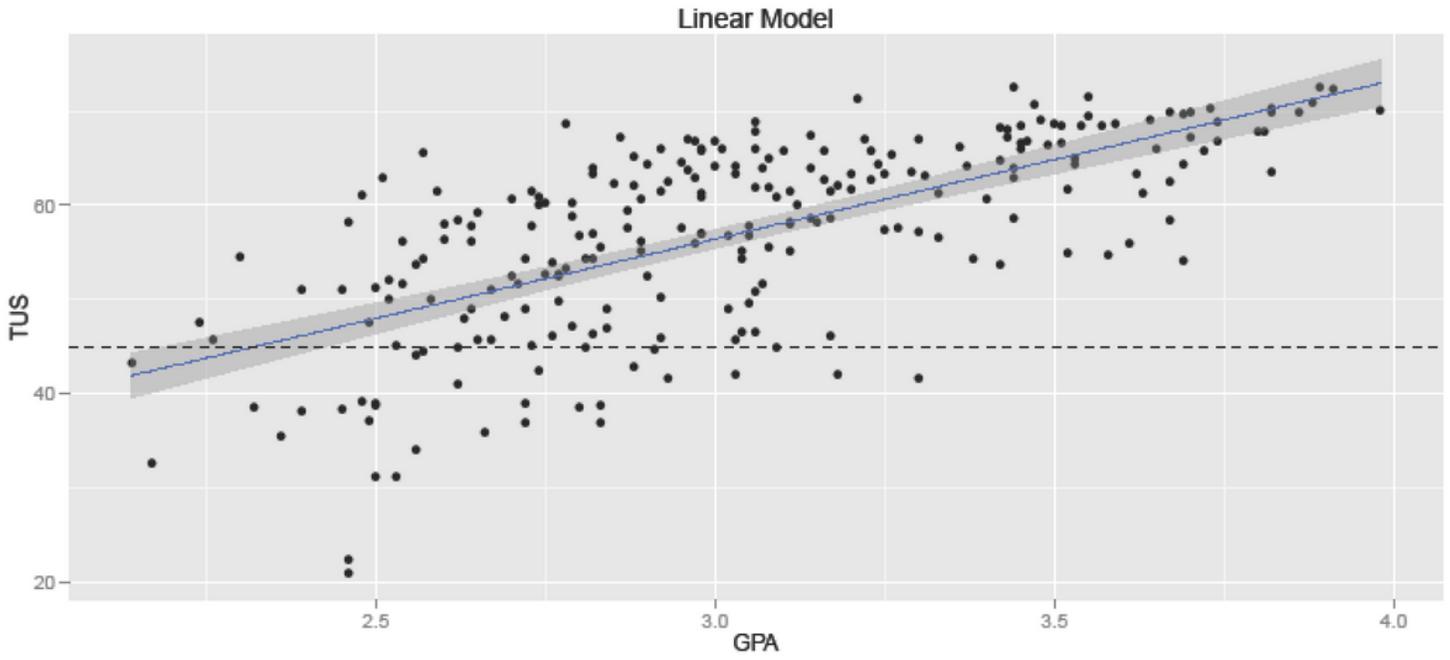


Figure 3

Final GPAs and TUS scores plot of all students with applied linear regression model