

UNIVERZITA KARLOVA V PRAZE

1. lékařská fakulta

DIZERTAČNÍ PRÁCE



Praha 2012

RNDr. Čestmír ŠTUKA, MBA



UNIVERZITA KARLOVA
V PRAZE



1. lékařská fakulta

ÚSPĚŠNOST STUDIA
Z POHLEDU MODERNÍCH METOD ANALÝZY DAT

STUDY SUCCESS
IN VIEW OF MODERN DATA ANALYZING METHOD

RNDr. Čestmír ŠTUKA, MBA

Školitel: Prof. MUDr. Štěpán Svačina, DrSc.

Praha 2012

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem řádně uvedl a citoval všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

V Praze, 16. 05. 2012

Čestmír Štuka

Identifikační záznam:

ŠTUKA, Čestmír. *Úspěšnost studia z pohledu moderních metod analýzy dat*. Praha, 2012. 86 s.
Dizertační práce. Univerzita Karlova, 1. lékařská fakulta v Praze. Vedoucí práce Prof. MUDr. Štěpán Svačina, DrSc.

Jméno a příjmení autora:	Čestmír Štuka
Název disertační práce:	Úspěšnost studia z pohledu moderních metod analýzy dat
Název disertační práce anglicky:	Study success in view of modern data analyzing method
Studijní program:	Biomedicínská informatika
Školitel:	Prof. MUDr. Štěpán Svačina, DrSc.
Rok obhajoby:	2012
Klíčová slova v češtině:	přijímací řízení, studijní úspěch, přijímací testy, prediktivní validita
Klíčová slova v angličtině:	admissions, academic achievement, entrance tests, predictive validity

Abstrakt:

Cílem práce je pomocí statistické analýzy údajů dostupných o uchazečích v době přijímacího řízení přispět k co nejlepší predikci budoucí studijní úspěšnosti studentů na 1. LF UK v Praze a umožnit tak kvalitnější výběr uchazečů.

Pro výběr uchazečů o studium bývají využívány informace o studijním výkonu na střední škole, přijímací testy, pohovory a další metody. V práci analyzujeme kritéria pro přijímání studentů na 1. lékařskou fakultu. Kromě tradiční formy přijímacích testů z profilových předmětů (B, CH, F) byla v roce 2002 otevřena možnost přijetí části studentů na základě jejich výsledků na střední škole.

Analyzujeme nejprve data 383 studentů přijatých ke studiu všeobecného lékařství v roce 1999, kdy byly naposledy pro všechny uchazeče známy i středoškolské studijní výkony a přitom všichni absolvovali přijímací test. Všichni studenti z této kohorty již dostudovali: 163 studium zanechalo a 220 studium úspěšně dokončilo. Použitím Pearsonových korelací mezi prediktory, shlukové analýzy a logistické regrese ukazujeme, že průměrný prospěch na střední škole umožňuje předpovídat celkový úspěch studia medicíny s podobnou přesností jako samotné přijímací testy, ale každý z těchto faktorů popisuje jinou dimenzi schopností uchazeče. Prokazujeme, že přijímání části studentů na základě jejich výtečného středoškolského prospěchu je racionální a opodstatněné. Ukazujeme, že současné využití obou zdrojů informací o uchazeči by mohlo zvýšit kvalitu výběru studentů.

V letech 2008-2010 byl ke standardním testům z profilových předmětů (B, CH, F) přidán test všeobecné studijní připravenosti (VSP). Zabýváme se proto porovnáním predikční validity jednotlivých testů pro předpověď studijních výsledků v prvních letech studia. Na datech z let 2002-2009 studujeme predikční validitu odborných testů, porovnáváme ji s predikční validitou testu VSP, abychom zjistili, jaký je přínos testu VSP k tradičním odborným testům. Ukazuje se, že testy z profilových předmětů prokazatelně přispívají ke správnému výběru studentů, zatím co příspěvek testu VSP je sporný.

Klíčová slova:

přijímací řízení, studijní úspěch, přijímací testy, prediktivní validita

Dissertation abstract:

The aim of the project is to find, by statistical analysis of the data available at the time of entrance examination, parameters that predict the student's ability to finish his/her studies successfully.

We analysed admission criteria based on admission test (AT) and on undergraduate grade-point average (uGPA) supplemented by some other criteria.

The study analyses admission procedure at the largest school of medicine in the Czech Republic. It is based on 1999 data of 383 students admitted that year. By use of Pearson's correlations between predictors, cluster analysis and logistic regression we show that preadmission grades predict the overall success in medical study with the same accuracy as admission tests but each of them describe different dimension of students' abilities. Simultaneous use of GPA and AT in admission process of all students should bring higher quality of selection process and also practical advantages for future analyses. At the same time we can confirm, the idea to accept part of students only on base an excellent GPA, could be considered a quite reasonable one.

General aptitude test (GAT) was temporarily added to admission process at the First Faculty of Medicine of Charles University in Prague, to prove predictive validity of GAT and to compare it with validity of current scientific knowledge test (ST). The added value of GAT was also examined. Prediction validity was estimated by Pearson correlation within cohorts (years 2002 – 2009) and by the correlation after each variable has been group-mean centred for all cohorts together. Incremental validity was assessed by testing submodels in linear and logistic regression models.

Although the predictive validity of GAT is significantly nonzero; it is markedly lower than predictive validity of ST. Also, scientific knowledge tests contribute significantly to the proper selection of students, whereas the contribution of GAT can be inferred only in part.

Keywords: admissions, academic achievement, entrance tests, predictive validity

Poděkování:

Chtěl bych tímto poděkovat Prof. MUDr. Štěpánu Svačinovi, DrSc. za inspiraci a vytrvalé směřování. Děkuji Patřicii Martinkové a Doc. Karlu Zvárovi za neocenitelné konzultace v průběhu celé práce a ing. Haně Mezulianikové za opakovanou pomoc při sběru dat.

Obsah

Úvod	10
Stručná historie přijímání studentů medicíny na (1. LF) UK.....	11
Přehled přijímacích řízení na českých lékařských fakultách	12
Přijímací řízení na 1. LF UK v Praze	13
Přijímací testy	13
Testové otázky	14
Přijímání studentů na základě výsledků ze střední školy.....	17
Problematika zápisu.....	17
Test všeobecných studijních předpokladů.....	18
Přijímací řízení na dalších lékařských fakultách v Čechách	20
2. lékařská fakulta Univerzity Karlovy v Praze	20
3. lékařská fakulta Univerzity Karlovy v Praze	20
Lékařská fakulta Univerzity Karlovy v Hradci Králové	21
Lékařská fakulta Univerzity Karlovy v Plzni	22
Lékařská fakulta Univerzity Palackého v Olomouci	22
Lékařská fakulta Masarykovy univerzity v Brně.....	23
Přehled přijímacích postupů na lékařské fakulty v zahraničí.....	24
Abecední seznam termínů z oboru testování a přijímacích řízení.....	24
Metody výběr uchazečů o studium v zahraničí.....	28
Ukažme proto na příkladech jak se tato problematika ve vyspělých zemích řeší.Spojené státy americké.....	29
SAT	30
Předmětové zkoušky SAT.....	35
Předběžný test SAT	35
ACT	35
Použití testů a středoškolského prospěchu pro výběr studentů na vysoké školy	38
Velká Británie	39
Austrálie	41
Izrael	42
PET	43
Problémy PET	44
Shrnutí literární rešerše	45
Předběžná studie.....	46
Studium prediktorů úspěšného absolvování studia na 1. LF UK	46
Cíle	46
Datový soubor.....	46
Popisné charakteristiky.....	47
Závěr	52
The prediction and probability for successful completion in medical study based on tests and pre-admission grades.....	53
ABSTRACT	53
INTRODUCTION	54
METHODS	56
Subjects.....	56
Admission conditions.....	56
Measures in the data	57
Statistical analysis	58
RESULTS.....	59
Relationship between different GPA and AT predictors.....	60

Optimal model for predicting study success	61
Current admission criteria	63
DISCUSSION	64
CONCLUSION	67
References:.....	68
Predictive Validity of Scientific Knowledge Test and Intellectual Aptitude Test Used in Admission for Medical School Performance during the First Year – Case Study	70
Abstract	70
Introduction.....	71
Admission procedures at Czech faculties of medicine	71
Admission procedure at the First Faculty of Medicine in Prague.....	72
Aims.....	73
Methods	74
Admission tests	74
Academic performance.....	75
Data set	77
Statistical analysis of data	78
Results	79
Discussion	81
References.....	82
Závěr.....	83
Literatura.....	84

Úvod

Cílem práce je přispět pomocí statistické analýzy údajů dostupných o uchazečích v době přijímacího řízení k lepší predikci budoucí studijní úspěšnosti studentů všeobecného lékařství na 1. LF UK v Praze.

Správný výběr uchazečů o studium medicíny je důležitý pro lékařskou fakultu i pro uchazeče. Účelem přijímacího řízení je vybrat uchazeče tak, aby byli schopni ve vymezeném čase úspěšně absolvovat studium. Případná nedokonalost výběru vede na jedné straně k neefektivnímu využití výukové kapacity školy, na druhé straně ke ztrátě času studentů, věnovaného nevhodně zvolenému oboru studia. Správně nastavené přijímací řízení by mělo pomoci studentům, aby zjistili, zda pro náročné studium mají příslušné vlohy.

Pro výběr studentů bývají využívány informace o studijním výkonu na střední škole, přijímací testy, pohovory a další metody.

Z hlediska kvality přijímacího řízení nás zajímají tři okruhy otázek:

- Metody přijímacího řízení pro výběr studentů
- Hledání optimálního prediktoru studijního úspěchu na základě přijímacích testů a výsledků ze střední školy
- Porovnání přínosu znalostních a schopnostních testů na 1. LF UK

Práce je členěna na tři části.

1. V první části podáváme přehled metod přijímacího řízení používaných na 1. lékařské fakultě, na dalších lékařských fakultách v Čechách a na vysokých školách zahraničí a zabýváme předběžnou studií zaměřenou na možnost využití dostupných informací pro predikci studijního úspěchu na lékařské fakultě.
2. Ve druhé části se formou publikace zabýváme vzájemnými vztahy prediktorů odvozených z přijímacích testů a středoškolských studijních výkonů.
3. Ve čtvrté části podáváme formou publikace srovnání predikční validity testu všeobecné studijní připravenosti se znalostními testy z profilových předmětů.

Stručná historie přijímání studentů medicíny na (1. LF) UK

- 1348 – V prvních 600 let po založení Karlovy Univerzity byli studenti přijímáni na základě postavení a majetku a to nejdříve na Artistickou (filozofickou) fakultu ve věku 14-16 let. Předpokládala se u nich jen určitá znalost čtení, psaní a počítání a především základy latiny, jakožto jazyku, v němž se v Evropě vzdělání šířilo. Artistickou fakultu úspěšně dokončila zhruba čtvrtina až třetina přijatých studentů. Teprve po absolvování této (z našeho pohledu spíše střední) školy mohl student pokračovat na vyšších fakultách - lékařské, právnické a teologické. Činila tak ovšem jen velmi malá část absolventů bakalářského stupně na Artistické fakultě, neboť navazující studium bylo podstatně náročnější a to i finančně. V pozdějších dobách představovala filozofická fakulta mezistupeň mezi gymnázií a nejvyšším vzděláním (Přenosil 2004).
- 1849 – v tomto roce byla filozofická fakulta zrovnoprávněna s ostatními fakultami a studenti se na ně mohli hlásit po absolvování gymnázia nebo rovnocenné střední školy a složení maturitní zkoušky.
- 1948 – komunistický režim zrušil dosavadní postupy a zavedl dvoukolové přijímací řízení skládající se z písemných testů a ústního pohovoru, který měl původně prověřit nejen vědomosti a schopnosti uchazečů, ale především jejich „kádrové předpoklady“ pro studium. Jak režim postupně erodoval, byla ideologická hlediska postupně nahrazována spíše formálními požadavky. V „plánovaném hospodářství“ socialistický stát zavedl i pevně stanovené kvóty pro přijímání studentů.
- 1990 – po změně režimu bylo na 1. LF UK v Praze zavedeno přijímání studentů na základě písemných testů ze tří profilových předmětů (Fyzika, Biologie, Chemie). Otázky používané v těchto znalostních testech vznikají na odborných ústavech 1. LF UK a jsou předem zveřejňovány, aby byla zajištěna rovnost přístupu ke studijním materiálům.
- 2002 – byla otevřena možnost přijetí na základě vynikajících výsledků ze střední školy.
- 2008 – byl k testům z profilových předmětů experimentálně připojen test „všeobecných studijních předpokladů“ (VSP), zaměřený na intelektuální schopnosti, a to zejména na logické a analytické schopnosti a prostorovou představivost.
- 2011 – Po analýze přínosů a zkušeností byl test VSP z přijímacích testů na 1. LF UK odebrán.

Přehled přijímacích řízení na českých lékařských fakultách

Pro výběr uchazečů se na všech sedmi českých lékařských fakultách využívají testy z fyziky, chemie a biologie. Na některých školách jsou pak testy doplněny pohovorem, interpretací textu, nebo testem „všeobecných studijních předpokladů“.

Od roku 2002 je na 1. LF UK v Praze a několika dalších českých lékařských fakultách zavedena možnost přijetí části uchazečů bez absolvování přijímacích testů, na základě průměrů známek na střední škole.

Přehledně jsou informace o přijímacím řízení pro rok 2010 na všech sedmi LF v ČR podány v Tabulce 1. A podrobněji pak v textech k jednotlivým fakultám.

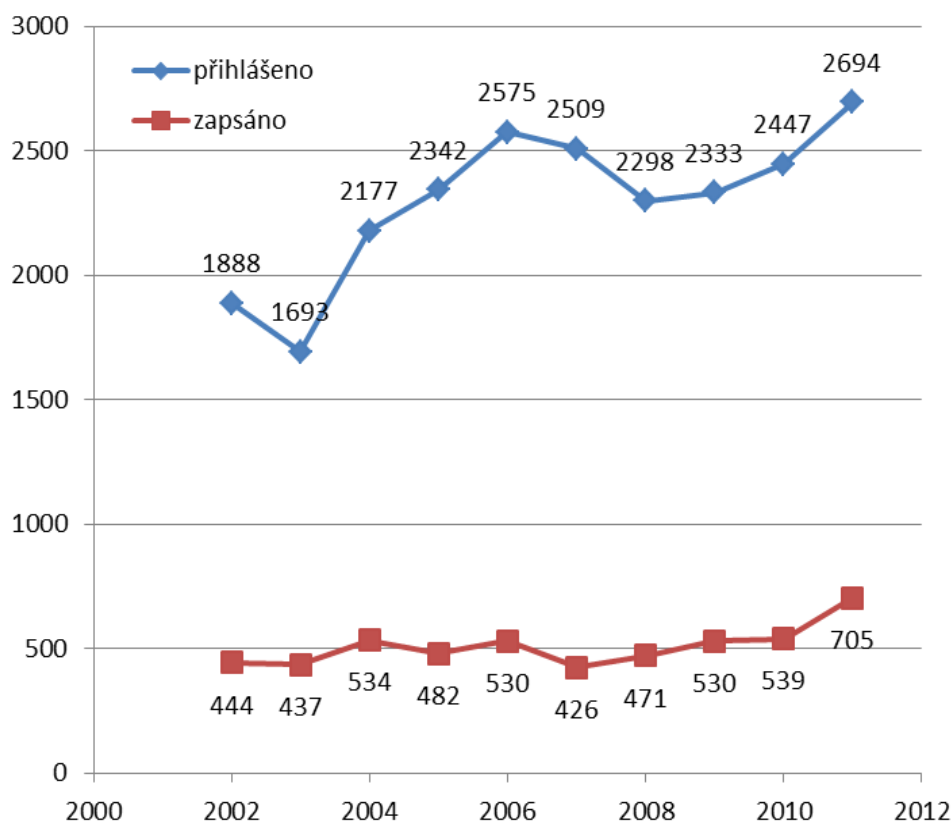
Tabulka 1. Přijímací řízení na lékařské fakulty v České republice (Topičová 2011)

Lékařská fakulta:	Test B, Ch, F	typové otázky zveřejněny	Test VSP	Pohovor	Interpretace textu	Možnost přijetí podle výsledků na SŠ	Počet přijatých *	Počet uchazečů
1. lékařská fakulta Univerzity Karlovy v Praze	1	1	1	0	0	1	450	3374
2. lékařská fakulta Univerzity Karlovy v Praze	1	1	1	1	0	0	150	1813
3. lékařská fakulta Univerzity Karlovy v Praze	1	0	0	1	1	1	155	1413
Lékařská fakulta Univerzity Karlovy v Hradci Králové	1	1	0	0	0	1	165	1415
Lékařská fakulta Univerzity Karlovy v Plzni	1	1	0	0	0	1	230	1850
Lékařská fakulta Univerzity Palackého v Olomouci	1	1	0	0	0	1	220	1838
Lékařská fakulta Masarykovy univerzity v Brně	1	1	0	0	0	0	400	2894

* počet přijatých je orientační, týká se studentů oboru všeobecného lékařství studovaného v národním jazyce přijímaných v roce 2010, kteří byli přijati, zapsali se a skutečně nastoupili ke studiu.

Přijímací řízení na 1. LF UK v Praze

1. LF UK v Praze je největší ze sedmi českých lékařských fakult. Na studium všeobecného lékařství v národním jazyce se každoročně запиše kolem 500 přijatých studentů z asi 3500 uchazečů.



Obr. 1. Počet uchazečů a studentů všeobecného lékařství v českém jazyce zapsaných v jednotlivých letech (Línová 2011).

Přijímací testy

V letním semestru akademického roku 1989 / 90 začalo pracovat nové kolegium děkana schválené akademickou radou. V tomto kolegiu zastával funkci proděkana pro studijní záležitosti doc. Petr Hach, který byl tehdejším děkanem prof. Pacovským pověřen vypracovat nový model přijímacího řízení, který by vyloučil námitky neúspěšných adeptů, že řízení je subjektivní a poplatné osobním náladám examinátorů.

Spolu s doc. Marešem a ve spolupráci s tehdejšími pracovníky kybernetického oddělení Fysiologického ústavu 1. lékařské fakulty sestavili model, který se v podstatě používá dodnes - tedy sada 3 testů s mnohočetným výběrem odpovědí, vyznačovaných na

odpovědním formuláři, který je po odevzdání zpracováván elektronicky. Výsledky testů se spočítají a až po úplném vyhodnocení spojí se jménem uchazeče. Identifikace záhlavních listů a testových formulářů je pak zajištěna čárovými kódy, které na odpovědní list a záhlaví list lepí učitel.

Pokud se v zadání některé otázky vyskytla chyba, pak byla otázka anulována u všech odevzdaných formulářů a příslušně sníženo bylo i dosažitelné maximum bodů.

Nový model umožnil zpracování výsledků přijímacích zkoušek během 12 hod. po posledním testu. Na jiných fakultách přitom trvalo tou dobou zpracování zkoušek i více než týden. Nový postup zcela odstranil námitky proti neobjektivnímu přijímacímu řízení. Praktickým výsledkem pak je, že po celou dobu, kdy fakulta tento model praktikuje, neuznal děkan v autoremeduře, ani rektor v přezkumném řízení jediné odvolání pro neobjektivitu přijímacích zkoušek (Hach 2010).

Formální postupy a principy přijímacího řízení upravuje Řád přijímacího řízení 1. lékařské fakulty Univerzity Karlovy v Praze (Zima 2008).

Testové otázky

Otázky používané ve znalostních testech vznikají na odborných ústavech 1. LF UK. U některých otázek může existovat více než jedna správná odpověď a jako dobré zodpovězení otázky oceněné jedním bodem se bere pouze zaškrtnutí všech správných jednotlivých odpovědí.

Pro každý ze tří předmětů existuje sada cca 1300 otázek, které jsou každoročně zhruba z jedné třetiny inovovány. Tyto otázky mají autorem přidělenou obtížnost v rozsahu 1-3.

Z těchto otázek je těsně před vlastním přijímacím řízením metodou náhodného výběru počítačem vybrán test (100 otázek) o požadovaném průměrném stupni obtížnosti (obvykle 2,5)¹. Vytiskne se jedno paré od každé varianty a automaticky se vytvoří datový nosič se správnými odpověďmi a oboje se okamžitě pečeti do obálky. Tento postup je u každého

¹ Vzhledem k tomu, že otázek s obtížností 3 je omezené množství a současně požaduje fakulta průměrnou obtížnost otázek 2,5, dochází v praxi k tomu, že obtížné otázky jsou při strojovém generování tří sad testů využity všechny (snižuje se náhodnost výběru) a přitom stále ještě není požadované průměrné obtížnosti dosaženo (praxi bývá průměrná obtížnost kolem 2,2).

předmětu zopakován třikrát. Jsou tedy připravena vlastně tři přijímací řízení. Tyto tři jinak než předmětem neoznačené obálky jsou uchovány v trezoru fakulty.

Den před přijímacím řízením děkan vylosuje jednu obálku pro každý předmět a z tam uložených matric jsou mimo fakultu rozmnoženy příslušné testy. Tisk probíhá za dozoru zástupců vedení fakulty tak, aby bylo zabezpečeno skartování chybných, nebo navíc vytištěných testů. Testy jsou okamžitě rozdělovány pro jednotlivé posluchárny a balíky s testy jsou ještě v rozmnožovně zapečetěny. Datové nosiče se správnými odpověďmi se odpečetují až po napsání příslušného testu v souvislosti s vyhodnocováním.

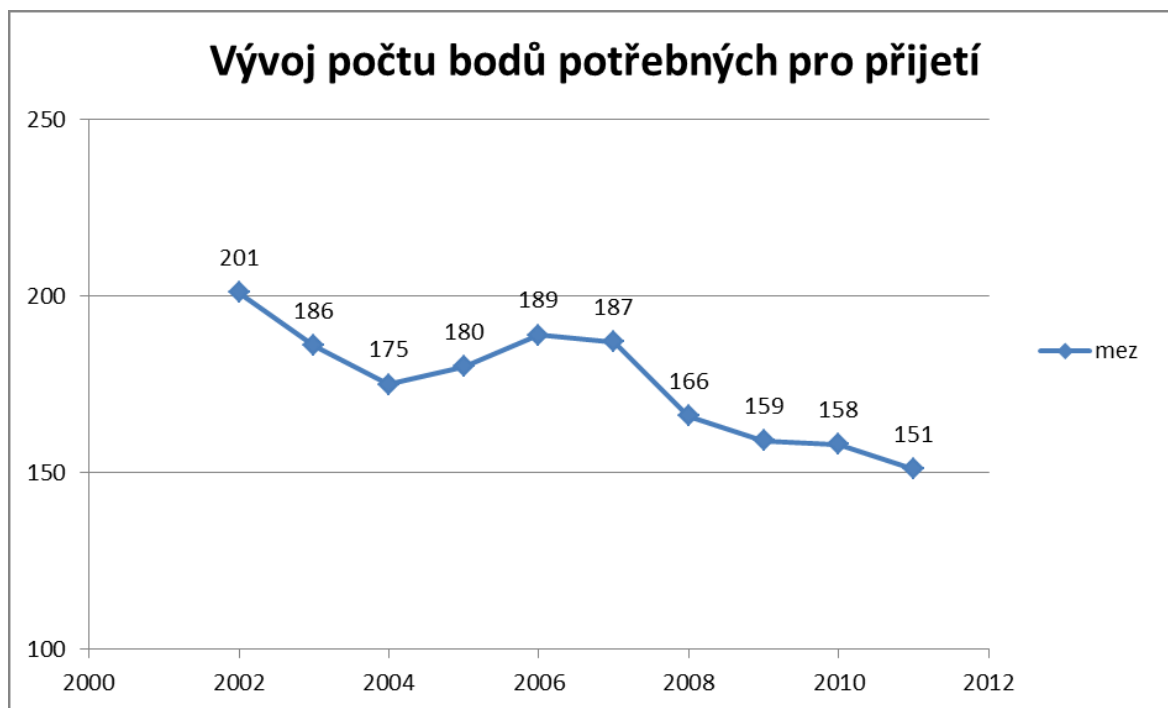
Anonymita testu po celou dobu zpracování je zaručena tím, že uchazeč odevzdává dva listy - osobní list se jménem a rodným číslem uchazeče a odpovědní list, který není nijak signován. Při odevzdávání nalepí učitel na oba listy identický čarový kód. Oba listy se oddělí a od té chvíle jsou zpracovávány nezávisle. Odpovědní listy jsou vyhodnocovány automaticky pomocí skenerů a programu a na jiném pracovišti jsou spojována rodná čísla a jména uchazečů s náhodným číslem na osobním listu.

Teprve po vyhodnocení všech odpovědních listů a přečtení všech identifikačních kódů dojde ke spojení odpovědí se jmény studentů. Po proběhnutí testů ze všech předmětů vytvoří program pořadí uchazečů podle počtu dosažených bodů.

Děkan pak, podle dosaženého pořadí, nabídne přijetí na fakultu takovému počtu uchazečů, aby byl naplněn počet studentů odpovídající kapacitě školy a schválený vědeckou radou a senátem fakulty.

Ve sledovaném období 1990-2010 jsme vynesli do grafu mez počtu bodů potřebných pro přijetí na studium všeobecného lékařství (obr. 2.). Počet bodů je normován na standardní počet bodů 300, aby se vyloučil vliv drobných odchylek v celkovém počtu hodnocených otázek v letech, kdy byla některá otázka dodatečně z testu vyloučena.

Přepočtená mez počtu bodů má ve sledovaném období zřetelně sestupnou tendenci. Interpretace toto faktu je nejasná. Může jít jak o posun v obtížnosti testu, tak o vývoj v připravenosti či kvalitě uchazečů o studium. Pro vyloučení posunu obtížnosti testu jsou v některých zemích (USA, Izrael, ...) v testech používány „normativní“ otázky (20-33% z celkového množství), které se ve dvojici po sobě jdoucích testů opakují. Slouží jednak k provázání dvojic testů mezi sebou, jednak umožňují testovat otázky samotné před jejich zařazením do vyhodnocované části testu.



Obr. 2. Vývoj počtu bodů potřebných pro přijetí na 1. LF UK v letech 2002-2011

Typové otázky jsou předem zveřejňovány v podobě tištěných skript. Zveřejněné otázky jsou podobné otázkám pro přijímací zkoušky, ale nejsou totožné. Mohou mít změněné pořadí odpovědí, odlišné numerické hodnoty v příkladech a podobně. Jednotlivá skripta obsahují cca 1 000 otázek pro každý předmět. Skripta je možné zakoupit na děkanátu fakulty.

Příprava testových otázek na ústavech 1. LF UK Praha je pokládána za velmi pracnou, nepopulární a málo hodnocenou. Životní cyklus otázek, tak jak je dnes nastaven, nepředpokládá a ani neumožňuje využít objektivní systém kontroly pro zlepšení kvality testových otázek (např. metodou položkové analýzy), protože otázky se z roku na rok mění a starší, u nichž by bylo možné validitu (po sedmi letech) objektivně stanovit se znovu již nepoužijí.

Přitom podle našich zjištění je validita i reliabilita otázek nevyrovnaná. Validita se pohybuje od otázek vysloveně přínosných až po otázky, které s celkovým studijním úspěchem korelují negativně. Tj. jejich správné zodpovězení indikuje studenty ve studiu posléze spíše neúspěšné.

Testové otázky podle našich vědomostí neprocházejí nezávislou vnitřní oponenturou a podle sdělení některých pedagogů je jejich zaměření často marginální, místo aby postihovaly podstatné problémy oboru (Vejsražka 2011).

Přijímání studentů na základě výsledků ze střední školy

V roce 2001 jsme na žádost vedení fakulty vypracovali studii „Statistické vyhodnocení přijímacího řízení a výsledků ze střední školy ve vztahu k úspěšnosti studia na vysoké škole“ (Štuka and Klaschka 2003). Ukázalo se, že známky ze střední školy významně korelují se schopností studentů dostudovat školu v řádném termínu.

Na základě této studie byla v roce 2002 na 1. LF UK v Praze otevřena možnost přijetí na základě výsledků ze střední školy. V následujících letech zavedly tuto možnost i další lékařské fakulty.

Podmínky přijetí:

1. Uchazeč musí ve školním roce, kdy žádá o přijetí úspěšně složit maturitní zkoušku na střední škole, přičemž jedním z předmětů zkoušky musí být český jazyk.
2. Výsledný průměr všech známek uvedených na přihlášce ke studiu není vyšší než 1,20 (nezapočítává se maturitní vysvědčení).
3. Uchazeč absolvoval výuku tří profilových předmětů fyzika, chemie, biologie v délce alespoň jeden rok v průběhu posledních dvou let studia a nebyl klasifikován horší známkou než 2 (chvalitebně). Podle kritéria uvedeného v bodě 2 se vytváří pořadí uchazečů o studium, kteří jsou pak pozváni k zápisu ke studiu.

Problematika zápisu

1. LF UK přijímá od roku 2002 část studentů bez přijímacích zkoušek, na základě studijního prospěchu na střední škole a několika omezujících podmínek, které vymezují typ školy a další požadované parametry. Protože možnost přijetí na základě výsledků na střední škole zavedly i další české lékařské fakulty za podobných podmínek jako 1. LF UK a studenti mohou podat více žádostí o přijetí, stává se, že student takto přijatý na jednu fakultu je současně přijat i na jiné lékařské fakulty. V analýze (Martínková, Zvára et al. 2009), kterou jsme provedli, se ukazuje, že významná část studentů, kterým bylo na základě vynikajících výsledků na střední škole nabídnuto přijetí na 1. LF UK, se na fakultu nezapíše a dají přednost jiné lékařské fakultě a to tím spíše, čím lepší mají středoškolský prospěch. Vazba na SŠ prospěch může souviset s tím, že hranice SŠ průměru pro přijetí na jiné škole (viz např. HK) může být posazena výše a část studentů s lepšími průměry tak má víc voleb než studenti s průměry hrošími.

Studenti, kteří přijímací zkoušky absolvovali, se zapisují s vyšší pravděpodobností bez ohledu na jejich průměry na střední škole.

Test všeobecných studijních předpokladů

V roce 2008 byl na 1. LF UK v Praze ke třem testům z profilových předmětů připojen test „všeobecných studijních předpokladů“ (VSP). Zatímco první tři testy dohromady vlastně vytvářejí „znalostní test“, nový test VSP je testem schopnostním. V zahraničí se pro odlišení testovaných vlastností občas používá termín test „krystalické“ a „fluidní“ inteligence. Přitom se krystalickou inteligencí míní schopnost uplatnit nabyté znalosti při řešení problémů a fluidní inteligencí se míní schopnost abstraktního myšlení. Test byl vypracován na Filozofické fakultě Univerzity Karlovy v Praze. Na rozdíl od prvních tří testů zaměřených na znalosti z chemie, biologie a fyziky, nový test byl zaměřen na intelektuální schopnosti, a to zejména na logické a analytické schopnosti a prostorovou představivost. Testové otázky formou i obsahem odpovídají otázkám z IQ testů.

Na rozdílný příspěvek znalostních a dovednostních testů pro predikci studijního výkonu upozorňují v případě testu BMAT shodně McManus a Emery. V práci (Emery and Bell 2009) a (McManus, Ferguson et al. 2011) ukazují, že v případě Biomedical Admission Testu (BMAT), používaného v UK, jeho Sekce 2 zaměřená na Scientific Knowledge and Applications, koreluje se studijním průměrem v prvním a druhém ročníku zřetelně lépe než Sekce 1 zaměřená na Aptitude and Skills. V článku (Emery, Bell et al. 2011) Emery dále ukazuje, že pro odhad studijního průměru v prvním ročníku na VŠ přispívá Sekce 2 k predikci signifikantně, zatímco Sekce 1 nikoliv.

Rozdílné příspěvky znalostních a schopnostních testů byly diskutovány již dříve - např. (Hach, Kraml et al. 1994) v odpovědi na Andělův článek (Anděl, Höschl et al. 1993) poukazují na to, že zařazení psychologických testů do přijímacího řízení na lékařských fakultách může být problematické a některé školy (např. v článku zmiňovaná Lékařská fakulta v Calgary v Kanadě) od nich po předchozích zkušenostech upustily.

Vhodnost použití testu inteligence pro výběr uchazečů o studia medicíny v UK ve srovnání s vhodností použití výsledků státní maturity (A levels) diskutoval například McManus (McManus, Powis et al. 2005). Ten shrnuje svoje předchozí poznatky (McManus, Smithers et al. 2003) i závěry jiných studií s konstatováním, že státní maturity (A level) budoucí akademickou úspěšnost predikují, ale test intelektuálních schopností nikoli. Současně se zde konstatuje, že školy (v UK) zavádějí testy intelektuálních schopností bez důkazů o jejich vhodnosti, přesnosti či přidané hodnotě.

V našem případě podobně jako v Emery porovnááme vhodnost inteligenčního testu a testu znalostí. Při měření predikční validity testů se většinou porovnává výsledek testu se „studijním výkonem“ nejčastěji v prvních ročnících studia, dále s výsledky jednotlivých postupových testů, celkového studijního průměru či výsledkem státní zkoušky. Zatímco však schopnost přijímacích testů predikovat studijní výkon v prvním ročníku lze zjistit s ročním či dvouletým odstupem od přijímacího řízení, pro zjištění validity testů vzhledem k predikci celkového výsledku studia by bylo potřeba čekat dlouhou řadu let. V případě šestiletých magisterských studijních oborů (jako například v ČR) by potřebný časový interval byl až sedmiletý, což je z hlediska potřeby posouzení aktuálních změn přijímacích testů mnohdy nepřijatelné.

Proto se snažíme odhadnout kvalitu testů, tedy jejich schopnost vysvětlit celkový studijní výsledek studia na základě informací, které jsou o studentech dostupné již po prvních ročnících studia. Jako faktor úspěšnosti uvažujeme jednak průměrný studijní prospěch v prvním ročníku na 1. LF a dále informaci, zda student zanechal studia do dvou let.

Přijímací řízení na dalších lékařských fakultách v Čechách

2. lékařská fakulta Univerzity Karlovy v Praze

Přijímací zkouška na obor Všeobecné lékařství je dvoukolová.

První kolo přijímací zkoušky:

Písemné testy z biologie, fyziky, chemie (celkem 75 otázek, každá otázka hodnocena 2 body) a dále písemný test všeobecných předpokladů pro studium medicíny (25 otázek, každá hodnocena 3 body), který zkoumá schopnost logické úvahy, porozumění rozsáhlejšímu odbornému textu, schopnost odečítat údaje z grafů, prostorovou představivost a smysl pro postřehnutí detailů - absolvují všichni přihlášení uchazeči.

Druhé kolo přijímací zkoušky:

Ústní pohovor - absolvují uchazeči, kteří dosáhnou nejlepších výsledků v prvním kole; je možné získat maximálně 60 bodů. Koná se před tříčlennou komisí pedagogických pracovníků fakulty. Cílem této zkoušky je zjistit komunikační dovednost uchazeče, jeho schopnost obhájit vlastní názor a zájem o zvolený obor.

Součet bodů z obou kol je jediným určujícím faktorem pro rozhodnutí o přijetí ke studiu.

Ke studiu jsou přijati uchazeči, kteří se podle počtu získaných bodů umístí na 1. až 150. místě a dále všichni, kteří dosáhnou stejného počtu bodů jako uchazeč na místě 150 (Havlová 2010).

3. lékařská fakulta Univerzity Karlovy v Praze

Přijímací zkouška je dvoukolová - má 2 části ve dvou dnech:

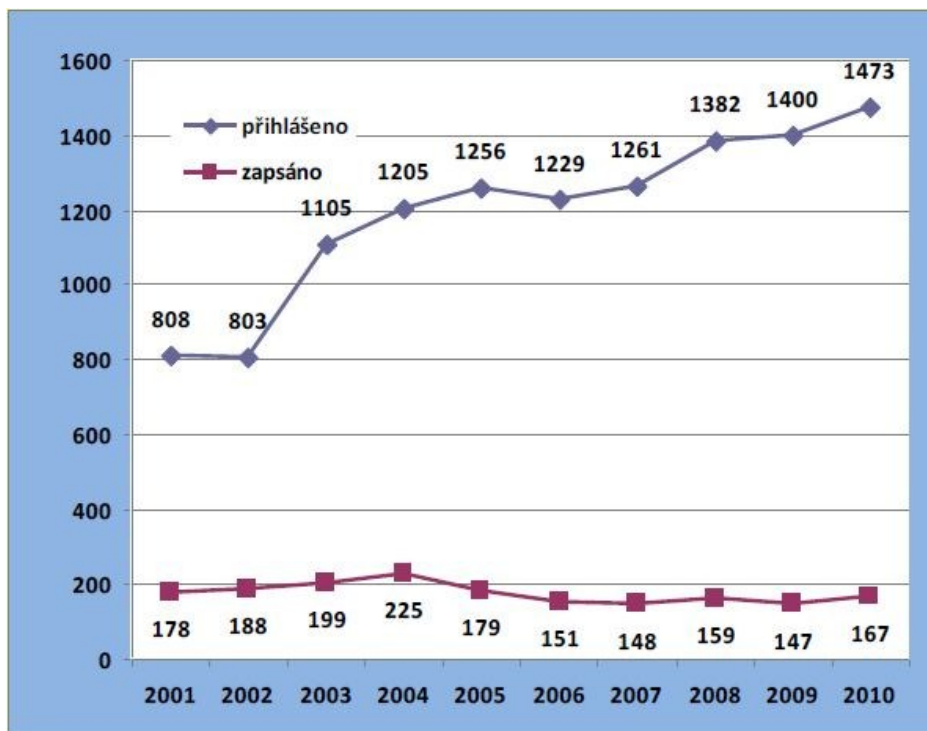
1. část – písemný test z biologie, fyziky a chemie

Celkem 90 otázek, z každého předmětu 30 otázek. Ze čtyř možných odpovědí je jen jedna správná odpověď. Minimální a maximální možný počet bodů za první část: 0 - 90

Podle dosažených výsledků postupuje do druhé části 300 nejlépe hodnocených uchazečů dle pořadí sestaveného podle dosažených výsledků.

2. část – ústní pohovor – zahrnuje interpretaci populárně vědeckého textu – max. 25 bodů, posouzení komunikačních schopností, schopnosti samostatného úsudku, všeobecného rozhledu – max. 37 bodů. (Pozn. druhá část není zkouškou znalostí).

Minimální a maximální možný počet bodů za druhou část je 12 – 62. Výsledkem přijímací zkoušky je kompozitní skór. Kompozitní skór se skládá ze součtu všech získaných bodů z přijímacího řízení, tj. z první a druhé části, kde hodnota průměru je korigovaná směrodatnou odchylkou celého souboru (Anděl 2010).



Obr. 3. Trend vývoje počtu přihlášených a zapsaných studentů na 3. LF UK za období 2001-2010 na magisterské studium "Všeobecné lékařství v českém jazyce" (Anděl 2010).

Lékařská fakulta Univerzity Karlovy v Hradci Králové

Student může být ke studiu přijat na základě přijímací zkoušky, nebo prospěchu na střední škole. Přijímací zkouška má formu testu z biologie, somatologie, fyziky a chemie, v jednom dnu pro všechny uchazeče. Test obsahuje 80 otázek, za které je možno získat 80 bodů. Odpovědi jsou hodnoceny: správná odpověď + 1 bod, žádná odpověď 0 bodů, nesprávná odpověď - 0,25 bodu. Další body je možno získat za účast v celostátních kolech olympiád (mat., chem., bio., fyz.).

Podmínky pro přijetí bez přijímací zkoušky byly:

- Absolvovat denní studium střední školy v České republice v aktuálním školním roce
- Průměr všech známek do 1,10 včetně,
- Absolvování tří ze čtyř profilových předmětů – biologie, fyzika, chemie, matematika, v délce alespoň jeden rok v posledních dvou letech studia (Slezák 2009).

Lékařská fakulta Univerzity Karlovy v Plzni

Přijímací řízení se skládá ze zkoušky z fyziky, chemie a biologie formou písemného testu a dosažení umístění v pořadí, které nepřesáhne předem stanovený počet přijatých (200). Rozsah požadovaných znalostí odpovídá učebním osnovám gymnázia. Fakulta vydává modelové otázky formou skript malého formátu. Otázky v přijímacím testu nemají formu otázek s mnohočetným výběrem odpovědi (multiple choice) a správně je vždy jen jedna odpověď. Otázky v testech nevycházejí přímo z modelových otázek. Test má po 50 otázkách z biologie, fyziky a chemie, každá správná odpověď je hodnocena jedním bodem a body se pak sčítají.

Lékařská fakulta Univerzity Palackého v Olomouci

Přijímací řízení je jednokolové a skládá se z testu z biologie, fyziky a chemie. Z každého předmětu je v testu 50 otázek. Z části jde o otázky mnohočetného výběru, přiřazovací otázky, otázky typu ANO-NE a příklady. Modelové otázky fakulta vydává a otázky v testech jsou prakticky identické. Na vypracování celého testu jsou 4 hodiny.

Podmínky pro přijetí bez vykonání přijímacích zkoušek:

Uchazeč má v daném studijním roce úspěšně složenou maturitní zkoušku obsahující zkoušku z jazyka českého. Celkový průměrný prospěch ze všech předmětů za celou dobu studia je nižší nebo roven 1,20. Výsledek maturity se do celkového průměru nezapočítává.

Uchazeč absolvoval výuku všech tří profilových předmětů fyzika, chemie, biologie v délce alespoň jednoho roku v průběhu posledních dvou let studia a nebyl klasifikován známkou horší než 2 (velmi dobře).

Lékařská fakulta Masarykovy univerzity v Brně

Zkoušky se konají formou písemných testů. Jsou testovány znalosti z odborných předmětů fyzika, chemie, biologie. Test všeobecné studijní připravenosti není po zkušenosti z let 2003-4 již dále součástí přijímacího řízení (Dušek 2004).

Testy se skládají ze 40 otázek s mnohočetným výběrem odpovědi. Délka trvání každého testu je 70 minut. Za každou správnou odpověď získá uchazeč jeden bod. Modelové otázky k přijímací zkoušce je možno zakoupit v prodejně knih LF MU.

Přehled přijímacích postupů na lékařské fakulty v zahraničí

Tato část práce podává přehled o metodách přijímacího řízení na zahraničních lékařských fakultách. Přehled je nutně selektivní a zaměřuje se především na ty země, kde je výběr studentů v ohnisku pozornosti a kde jsou tyto otázky často diskutovány v odborném tisku. Pro orientaci v používaných metodách výběru studentů přinášíme seznam často se vyskytujících zkratk v zahraničních odborných textech.

Abecední seznam termínů z oboru testování a přijímacích řízení

A-level

A-level je označení pro státní maturity ve Velké Británii. Pro výběr studentů na lékařské fakulty bývala požadována nejvyšší hodnocení ze tří předmětů (A), které musí obsahovat chemii a alespoň jednu další přírodní vědu nebo matematiku. Stupnice je A, B, C, D, E. Protože nejlepší lékařské fakulty mají tradičně několikanásobek uchazečů o místa, docházelo k tomu, že téměř všichni uchazeči měli nejlepší hodnocení - tři A ve státních maturitách. Takové hodnocení neposkytuje pro malou variabilitu vhodné podklady pro výběr studentů (McManus, Smithers et al. 2003).

ACT

American College Testing – je jeden ze dvou nejpoužívanějších přijímacích testů na vysoké školy v USA a v Kanadě. Většina škol dává studentům na výběr, který z testů budou absolvovat a stanovují jen počty bodů nutné pro přijetí v tom kterém testu. ACT se skládá se ze čtyř částí: Angličtiny (45 min.), matematiky (60 min.), čtení (35 min.) a vědeckého myšlení (35 min.).

BMAT

BioMedical Admissions Test - test používaný ve Velké Británii na tradiční lékařské fakulty s rozdělením preklinické a klinické výuky a důrazem na výuku vědy v prvních letech studia. BMAT je následníkem Medical and Veterinary Admissions Test (MVAT). BMAT hodnotí potenciál kandidátů k akademicky náročnému studiu biomedicíny na vysokoškolské úrovni. Test je náročný, aby se efektivně diskriminoval mezi žadateli o studium na vysoké škole.

Úspěšné absolvování BMAT je podmínkou pro na Cambridge, na Imperial College v Londýně, na Oxford a University College London. Informace o BMAT lze nalézt na webových stránkách www.bmat.org.uk.

Class rank

Pořadí výkonu studenta střední školy ve srovnání s ostatními studenty ve třídě. Viz. též high school class rank - HSCR.

EMQs

EMQs nebo též EMI (extended matching question/items) jsou nově zaváděné „rozšířené otázky s mnohočetným výběrem“. Důvodem pro jejich zavedení je, že MCQ svou formou svádějí k biflování, nikoli k pochopení. Charakteristickými znaky EMQ jsou: Mnohočetné odpovědi – alespoň osm možných odpovědí reprezentujících nejméně dva různé scénáře. Otázka má být položena tak, aby student byl schopen zformulovat správnou odpověď, aniž by předtím viděl nabízené varianty. (Sanju 2003)

FYGPA

First Year Grade Point Average – studijní průměry v prvním ročníku vysoké školy jsou používány pro odhad akademické výkonnosti studenta na vysoké škole.

GAMSAT

Graduate Australian Medical School Admissions Test - je test pro výběr uchazečů o magisterské studium medicíny, stomatologie a veterinárního lékařství vyvinutý v roce 1995 Australskou radou pro výzkum vzdělávání (ACER). Používá se pro výběr studentů na magisterský stupeň studia zdravotnických vysokých škol v Austrálii a od roku 1999 i na některých školách ve Velké Británii a v Irsku. GAMSAT zahrnuje testování schopnosti kriticky myslet, analyzovat informace a vyjadřovat své myšlenky jasně a efektivně.

Úspěšné absolvování GAMSAT je podmínkou přijetí do magisterských programů na univerzitě St Georgeova Londýna, Universitě of Nottingham v Derby, Universitě of Wales Swansea a Keele University. Registrace ke GAMSAT je mezi červnem a srpnem, a zkouška se koná jednou za rok v polovině září. Další informace lze nalézt na webových stránkách GAMSAT www.gamsatuk.org.

GAT

General aptitude test – test všeobecné studijní připravenosti je souhrnný název pro schopnostní testy, které testují rozumové schopnost uchazečů (na rozdíl od testů znalostních). Testu VSP bývají zpravidla otázky zaměřené na prostorové vztahy, logické souvislosti, a představivost, velmi podobné otázkám z IQ testů.

GCSE

General Certificate of Secondary Education – certifikát o středoškolském vzdělání v příslušném oboru používaný v Anglii, srovnatelný s maturitním vysvědčením. Vydává se 14-16 ti letým studentům po splnění příslušné zkoušky. Proti používání této zkoušky se postavila významná Konfederace britského průmyslu (CBI) sdružující 240 tis. společností. (<http://www.guardian.co.uk/education/2012/may/23/cbi-call-gcse-exams>) Podle šetření CBI vede tato zkouška jen k biflování na zkoušku, místo k výuce prakticky použitelných znalostí. Její význam je sporný, když od roku 2014 má být školní docházka do 18 let (kdy se skládají A-levels) povinná. Proto nemá smysl v 16 absolvovat GCSE.

GPA

Grade point average – studijní průměry. Studijní průměry ze střední školy bývají jako dobrý prediktor úspěšnosti studia zařazovány mezi výběrová kritéria pro přijetí na lékařské fakulty (Didier, Kreiter et al. 2006).

HSCR

High School Class Rank je měřítko studijní výkonnosti konkrétního studenta v poměru k výkonu ostatních ve třídě. Jiné označení pro tento parametr je Clas rank. Vypočte se jako pořadí studenta stanovené na základě GPA a vydělené počtem studentů ve třídě. Výsledkem je percentil nejlepších studentů, mezi něž student patří. Toto pořadí poskytuje asi 45% středních škol. Velké veřejné školy poskytují tento údaj častěji, než malé soukromé školy. HSCR se spolu s GPA často používá pro ohodnocení studenta při přijímání na vysoké školy.

HSGPA

High-school grade point average – označení pro studijní průměr na střední škole (USA). Viz. též GPA a uGPA.

MCAT

Medical College Admissin Test – standardizovaný „počítačový“ přijímací test na lékařské fakulty v USA, zavedený Asociací amerických lékařských fakult v USA.

MCAT(R)

V letech 1991-2 byl MCAT znovu revidován a restrukturalizován a jeho nová podoba nese výše uvedené označení.

MCQs (MULTIPLE CHOICE TEST)

Testy s mnohočetným výběrem odpovědí jsou patrně nejběžnější formou neverbálního ověřování znalostí na lékařských fakultách (Kasal 1997). Správná metodika tvorby otázek je klíčová pro použitelnost testu a je jí věnována v literatuře velká pozornost (Albanese, Kent et al. 1977), (Albanese, Kent et al. 1979), (Downing 2002).

SAT

Standardized Admissions Tests – tvoří spolu s **ACT** dva nejrozšířenější testy pro zjišťování připravenosti středoškoláků na vysokou školu v USA. V aktuální podobě platné od roku 2005 trvá SAT tři a tři čtvrtě hodiny a skládá se ze tří částí: Kritického čtení, matematiky a psaní. Za každou část je možné dosáhnout až 800 bodů. V testu je zahrnuta „experimentální“ část, která se nepoužívá pro vyhodnocení studentových schopností, ale pro zhodnocení otázky samotné, pro případné budoucí použití v testech SAT.

SBAs

Single Best Answer Questions – nově zaváděný formát testových otázek s jednou správnou odpovědí. Používá se především v klinických letech lékařských vysokých škol, neboť umožňuje lépe testovat nejen „vědět“, ale i „vědět jak“. Výhody zavedení tohoto formátu při výuce radiologie popisuje například McCoubrie (McCoubrie and McKnight 2008)

uGPA

Undergraduate grade point average - průměr známek na střední škole často používaný pro predikci studijního úspěchu na lékařských fakultách. Viz. též GPA a HSGPA.

Prediktivní validitou uGPA pro zlepšení výběru uchazečů úspěšných v prvním ročníku se zabýval Hall. Ukazuje, že kombinace uGPA, MCAT skóre a informace o „kvalitě“

střední školy je velmi užitečná při identifikaci studentů, kteří byli úspěšní v prvním ročníku (Hall and Bailey 1992).

UKCAT

UK Clinical Aptitude Test – je test vytvořený v roce 2006 konsorciem britských lékařských a stomatologických fakult pro testování duševních schopností uchazečů. UKCAT je navržen tak, aby testoval vlastnosti, které jsou považovány za důležité pro zdravotnické pracovníky. Jeho cílem je zajistit, aby žadatelé měli vhodné duševní schopnosti, postoje a profesní chování pro budoucí práci lékaře. Cílem testu tedy není predikovat akademický úspěch, který je predikován dobře pomocí A-levels, GCSE, nebo GPA. Test je tedy zaměřen na schopnost kritického logického myšlení a schopnost vyvozovat závěry. UKCAT se používá jako součást výběrového řízení na 26 lékařských škol ve Velké Británii. Podrobné informace lze nalézt na webových stránkách UKCAT www.ukcat.ac.uk.

Přínos toto testu je předmětem diskuze. (Fernando, Prescott et al. 2009). Lynch vyhodnotil prediktivní validitu UKCAT vůči prvnímu ročníku jako zanedbatelnou (Lynch, MacKenzie et al. 2009). Přítomnost tohoto testu podněcuje v UK diskuzi o vhodnosti psychologického testování uchazečů pro výběr studentů medicíny.

UMAT

Undergraduate Medicine and Health Sciences Admission Test se používá pro výběr středoškolských uchazečů o studium medicíny v Austrálii a na Novém Zélandu. Po absolvování bakalářského stupně jsou uchazeči o navazující „magisterské“ studium vybíráni pomocí GAMSAT.

USMLE

United States Medical Licensing Examination – je oficiální zkouška pro absolventy lékařských fakult pro vstup do postgraduálních programů klinické medicíny (Coumarbatch, Robinson et al. 2010). Margolis se zabýval šetřením, zda je dobře nastaveno mezní skóre této zkoušky a přináší důkazy pro správnost jejího nastavení (Margolis, Clauser et al. 2010).

Metody výběr uchazečů o studium v zahraničí

V zemích, které přikládají objektivním metodám výběru studentů velký význam se během doby vyvinuly metody a postupy testování a výběru studentů, jejichž znalost může být cenná i v našem prostředí. Příkladem může být provázání jednotlivých ročníků testů SAT ve Spojených státech, které umožňuje meziroční porovnání výkonu studentů a obtížností jednotlivých složek testů.

Ukažme proto na příkladech jak se tato problematika ve vyspělých zemích řeší.

Spojené státy americké

Přesto, že průměry známek ze střední školy (GPA) jsou významným prediktorem vysokoškolských studijních výkonů (Kobrin, Patterson et al. 2008), velké rozdíly v klasifikaci mezi školami si vyžadují použití nástroje, který by hodnotil studenty bez ohledu na jejich střední školu. Nejznámějším a nejpoužívanějším přijímacím testem ve Spojených státech je SAT. Ročně jej absolvuje více než 2 miliony studentů. (collegeboard.org 2011).

Není to však jediný nástroj pro výběr studentů. Kromě průměrů známek ze střední školy (GPA) se od roku 1959 používá ještě konkurenční American College Test, kterému se budeme věnovat v pozdějších odstavcích.

SAT

Test rozumových schopností SAT (SAT Reasoning Test, původně Scholastic Aptitude Test a od roku 1994 pak Scholastic Assessment Test) je dnes pravděpodobně nejznámějším testem studijních předpokladů. Byl zaveden americkou neziskovou společností „College Board“ založenou v roce 1900 s cílem standardizace přijímacího procesu na univerzitách a vysokých školách ve Spojených státech (McDonald, Newton et al. 2001).

Testy byly původně zaměřeny na hodnocení konkrétních znalostí z angličtiny, matematiky, řečtiny, latiny, historie a chemie. Verze testů SAT vydaná v roce 1926 byla již podobná té současné a vycházela ze zkušenosti s testováním IQ u dvou milionů rekrutů během první světové války. V své dnešní podobě SAT hodnotí verbální a matematické schopnosti, a to zejména formou otázek s možností výběru z více odpovědí (multiple choice).

Za jeho přínos se považuje nejen to, že měří připravenost absolventů středních škol pro vysokoškolská studia s ověřenou prediktivní validitou vůči studijním průměrům v prvním ročníku vysokých škol (Kobrin, Patterson et al. 2008), ale že to dělá nezávisle na sociálních faktorech, jako jsou pohlaví či etnický původ studentů (Clare 1999).

Mnoho pozornosti proto bylo věnováno „spravedlivosti“ testu. Diskutovaly se náznaky, že slovní i matematická část poněkud straní mužům. Při podrobném zkoumání se ukázalo, že nelze vyloučit, že příčinou pozorované nesymetrie byl náročnější předvýběr mužských absolventů zkoušek SAT. Ještě méně přesvědčivé jsou důkazy pro etnickou nespravedlivost testu. Afroameričané sice dosahují skóre téměř o směrodatnou odchylku

nižší než bílí Američané, ale podobně tomu je i u jiných testů, které měří aspekty inteligence. Otázku spravedlnosti testů by pravděpodobně nejlépe šlo řešit přezkoumáním toho, jak přesně výsledky testů předpovídají akademický výkon pro různé skupiny.

Testy SAT byly vytvořeny, aby sjednotily přijímací postupy různých vysokých škol v USA. Ačkoli tyto testy svou úlohu v podstatě plní, po vlnách kritiky mnohé školy přehodnotily způsob jejich používání a používají je buď ve spojení s dalšími metodami výběru, nebo od nich zcela ustoupily.

Jistým rizikem pro používání SAT je paradoxně jeho rozšířenost. Na některých středních školách byly již zavedeny přípravné kurzy pro SAT pořádané na úkor středoškolské výuky. Existuje reálné riziko, že na výběrových středních školách se mohou projevit na studentech výsledky cíleného koučování ve zkreslení výsledků testu.

Diskutuje se rovněž možné riziko inflace vysokoškolského vzdělání, pokud by byl zrušen SAT a pokud by neexistoval národní testovací program. Nicméně na školách, kde SAT přestal být pro přijetí povinný k žádným problémům s kvalitou výuky nedošlo.

Struktura testu SAT

Od roku 2005 se SAT sestává ze tří částí – matematického uvažování, kritického čtení a psaní. SAT není primárně určen k testování logiky nebo abstraktního myšlení, nýbrž k ohodnocení znalostí a dovedností které mají studenti získat na střední škole.

- **Kritické čtení** (SAT-CR, 70 min.) je zaměřeno na čtení a porozumění textu a schopnost dokončit větu pomocí nabídnutých slov.
- **Matematika** (SAT M, 70 min.) testuje znalost aritmetických operací, algebry, geometrie, statistiky a pravděpodobnosti.
- **Psaní** (SAT W, 60 min.) zahrnuje krátký esej, odpovědi na otázky s mnohočetným výběrem odpovědí zaměřené na hledání chyb a znalost gramatiky. Tato část byla do SAT přidána jako poslední v březnu 2005.

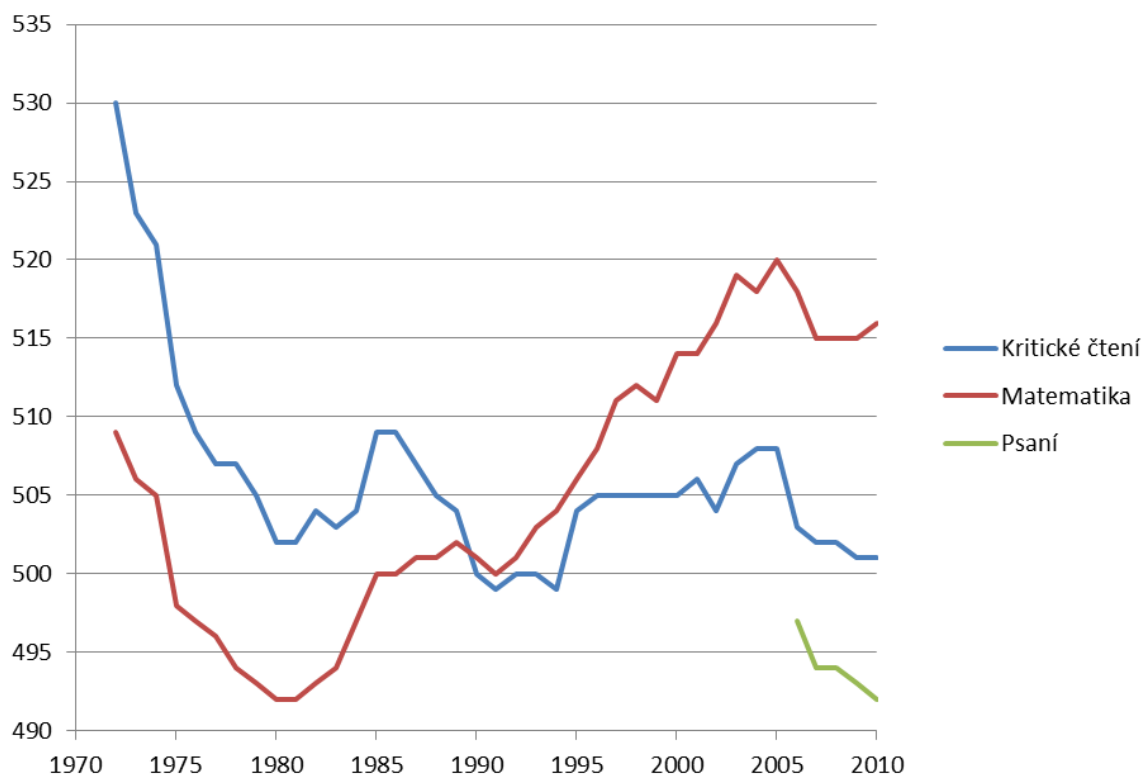
Tabulka 2.: Korelace jednotlivých prediktorů a jejich kombinací s FYGPA (Kobrin, Patterson et al. 2008)

Prediktor	Korelace s FYGPA
1. HSGPA	0,36
2. SAT-CR	0,29
3. SAT-M	0,26
4. SAT-W	0,33
5. SAT-M, SAT-CR	0,32
6. HSGPA, SAT-M, SAT-CR	0,44
7. SAT-CR, SAT-M, SAT-W	0,35
8. HSGPA, SAT-CR, SAT-M, SAT-W	0,46

Za každou část může uchazeč dostat od 200 do 800 bodů. Skóre má přibližně normální rozdělení s průměrem kolem 500 a směrodatnou odchylkou 110 bodů. Studenti obdrží informace o výsledku jednotlivých testů do 3 týdnů po absolvování zkoušky.

Během prvních let po zavedení SAT nebyly známy relace mezi jednotlivými verzemi testů. Aby byla možné vzájemně porovnat různé verze testů, používá se od roku 1941 v každém novém testu přibližně 20% otázek z testu předcházejícího.(Wainer 1999)

V roce 1995 byl test upraven tak, aby se medián opět nacházel na hodnotě 500 a směrodatná odchylka byla 100. Důvod předchozího poklesu není zcela zřejmý. Oficiální vysvětlení to označuje za důsledek rostoucí popularity SAT, který využívá stále více studentů i z méně výběrových škol. Dosahované výsledky mají dlouhodobě klesající tendenci dokonce i u nejlepších studentů. Za dvacet jedna let mezi roky 1972 a 1993 poklesl počet studentů, kteří dosáhli 600 a více bodů z „verbální“ části textu o 36% a to přesto, že počet zkoušených vzrostl. (Refor 1996)



Obr. 4. Vývoj středních hodnot počtů bodů z jednotlivých testů zkoušky SAT

(zdroj dat: Total Group Profile Report 2010,

<http://professionals.collegeboard.com/profdownload/cbs-2010-national-TOTAL-GROUP.pdf>)

Na grafu vývoje průměrných počtů bodů je nápadný trend rostoucího rozdílu mezi obtížností matematiky a kritického čtení. James Lech se věnoval tomuto jevu ve své dizertační práci z roku 2007 (Lech 2007). Navrhuje pro vysvětlení tohoto jevu několik hypotéz, z nichž nejznámější předpokládá, že je to projev eroze akademických znalostí, díky níž jsou kladené otázky z matematiky stále lehčí. Pro otestování této hypotézy byly vybrané otázky z matematiky rozeslány 1500 středoškolským učitelům, kteří měli stanovit obtížnost, aniž by věděli, odkud otázky pocházejí. Ukázalo se, že otázky jsou v průběhu času stále těžší.

Alternativní vysvětlení hledal Richard Rothstein, který v článku v New York Times (Rothstein 2002) v roce 2002 připustil, že skutečnou příčinu nikdo neví, ale že jednou z pravděpodobných příčin by mohlo být, že studenti se učí matematiku většinou ve škole, zatímco gramotnost si osvojují doma. Svou roli by v tom mohl hrát i nástup „televizní kultury“ vytěsňující postupně čtení z programu mladých Američanů.

Na jinou z možných příčin poukazuje Lech (Lech 2007), který připomíná strmý růst počtu studentů asijského původu ve sledovaných letech. Asijsí studenti přitom dosahují

výrazně vyšších skóre v matematické části SAT (580), než je celostátní průměr (520) a současně jejich jazykové znalosti jsou spíše průměrné.

Oba autoři souhlasně poukazují i na lepší kvalitu, organizaci a podporu učitelů matematiky oproti jejich humanitně zaměřeným kolegům.

Vztah SAT a IQ

Frey a Detterman (Frey and Detterman 2004) studovali vztah SAT a obecné inteligence (general intelligence **g**) a našli mezi nimi v obou konkrétních zkoumaných případech tak silnou korelaci, že zformulovali závěr, že SAT je navzdory ujištěním správců testu především testem inteligence. Dále poukazují na to, že výsledky SAT umožňují s uspokojivou přesností odhadnout IQ a přinášejí pro tento výpočet regresní rovnici. Tyto závěry zpřesnil a potvrdil Beaujean ve studii z roku 2006 (Beaujean, Firmin et al. 2006).

Problémy SAT

Zkoušky SAT jsou často napadány pro znevýhodňování některých etnik. V roce 1994 bylo z testů odstraněno hledání antonym (opaků) pro nejednoznačnost některých otázek položených touto formou. V roce 2005 byly po kritice University of California z verbální části testu odstraněny otázky na hledání analogií a z matematické části kvantitativní srovnávání.

Jedním z mediálně nejznámějších příkladů kulturně zaujatých otázek byla otázka na nalezení analogie podobné vztahu „běžec“-„maraton“ pro pojem „veslař“- („regata“). Správné zodpovězení této otázky předpokládá znalost kulturního prostředí bílé populace, v níže je veslování oblíbeným sportem. Otázku zodpovědělo správně 53% bílých uchazečů a jen 22% černých uchazečů.

V roce 2005 došlo k nesprávnému ohodnocení asi 4000 studentů v důsledku technické chyby při čtení dotazníků.

Na datech o studentech skládajících zkoušky SAT v roce 2009 bylo ukázáno, že průměrné skóre silně závisí na příjmu rodiny. Studenti z rodin s příjmem do 20 tis. USD měli průměrné skóre 1310, zatímco studentu z rodin s příjmem přes 200 tis. USD měli průměrné skóre 1715, tedy o 405 bodů více (DomeSAT 2009). Interpretace této korelace není jednotná. Vysvětluje se jak lepším vzděláním bohatších, tak dědičností inteligence.

Nahrává to však kritikům SAT, kteří tvrdí, že SAT zvýhodňuje bohaté. V reakci na tuto kritiku některé školy přestaly SAT pro přijetí vyžadovat.

Předmětové zkoušky SAT

SAT Subject Test (dříve SAT II) je název pro předmětově zaměřené testy SAT. Byly zavedeny na základě úspěchu SAT. Pomocí otázek s mnohočetným výběrem odpovědi testují znalosti ve dvaceti oborech, počínaje literaturou, přes americkou historii, světové dějiny, matematiku, biologii, chemii, fyziku a konče jazyky. Původně byl v SAT II i test psaní, který se pak od roku 2005 stal součástí SAT.

Na každý test má student hodinu času. Každý test je hodnocen na stupnici 200-800 bodů, ale některé testy byly nastaveny tak, že 200 bodů nebylo prakticky možné dostat.

Dva předmětové testy jsou obvykle vyžadovány pro přijetí na specificky zaměřených školách. Například stavební fakulty obvykle vyžadují dosažení jisté úrovně v testech z matematiky a chemie, nebo fyziky. Některé prestižní školy mohou vyžadovat (nebo silně doporučovat) předmětové zkoušky ze tří oborů.

Není nijak výjimečné, že škola místo testů SAT dává přednost alternativnímu testu ACT, ale k tomu vyžaduje předmětové testy SAT.

Předběžný test SAT

College Board rovněž nabízí Preliminary SAT (PSAT), který mohou studenti absolvovat ve druhém ročníku střední školy. Je velmi podobný samotnému SAT a má sloužit studentům k otestování silných a slabých stránek jejich středoškolského studia.

ACT

I když je test SAT ve Spojených státech zdaleka nejznámější a nejrozšířenější, není to test jediný. Jeho hlavním konkurentem je American College Test (ACT) zavedený roku 1959. Jedná se o standardizovaný test pro přijímání na americké vysoké školy produkovaný společností ACT, inc. Zkouška ACT sestává v základu ze čtyř samostatných testů z angličtiny, matematiky, čtení, logického myšlení a vědy. Testy jsou hodnoceny individuálně na stupnici 1-36 a souhrnné skóre je vypočítáno jako průměr z jednotlivých

testů. Ačkoli je ACT zaměřen odlišně než SAT a více zohledňuje středoškolské osnovy, jeho korelace se SAT dosahuje 0,9 (Schneider and Dorans 1999) naznačuje, že jsou testy v podstatě zástupné.

Na rozdíl od SAT není hodnoceno negativně, zaškrtně-li student v otázce s výběrem odpovědi i nesprávné varianty odpovědi.

ACT sestává z těchto subtestů:

- Test z angličtiny je 45 minutový a obsahuje 75 otázek zaměřených na míru porozumění psané angličtině, na interpunkci a standardní konvence a na rétorické dovednosti. Zkouška se skládá z pěti textových pasáží, po nichž následují otázky s mnohočetným výběrem odpovědi.
- Test z matematiky je 60 minutový a obsahuje 60 otázek. Je navržen tak, aby hodnotil dovednosti, které studenti zpravidla nabydou na střední škole, ale netestuje znalost složitých vzorců a výpočty. Zkouška zahrnuje otázky z algebry, geometrie, planimetrie a základů trigonometrie.
- Test ze čtení je 35 minutový. Ve čtyřiceti otázkách testuje porozumění psanému textu ve formě čtyř prozaických pasáží týkající se společenských věd, přírodních věd, humanitních věd a beletrie.
- Test vědeckého uvažování je 35 minutový a obsahuje 40 otázek s mnohočetným výběrem odpovědi. Test obsahuje sedm sad vstupních informací, které musí uživatel testu posoudit a odpovědět na přiložené otázky.

Protože mezi oba testy jsou v principu srovnatelné, používají se na některých vysokých školách pro porovnání převodní tabulky umocňující dobře odhadnout, jaký počet bodů by uchazeč získal v druhém testu, než který absolvoval. Vysoké školy zveřejňují obvykle požadované skóre pro oba testy a umožňují uchazečům, aby si test vybrali sami.

MCAT

Medical College Admissin Test – standardizovaný a od roku 2006 „počítačový“ přijímací test na lékařské fakulty v USA, zavedený Asociací amerických lékařských fakult v USA.

Počátek vývoje testu spadá do roku 1920, kdy byl zaznamenán významný pokles úspěšnosti studia medicíny z 95% na 50% a byla proto formulována potřeba studenty lépe vybírat před vstupem na vysokou školu. Test se skládá ze čtyř částí:

- Fyzikální vědy (PS)
- Slovní myšlení (VR)
- Psaní textů (WAS)
- Biologické vědy (BS)

Prediktivní validitou části těchto testů (Writing Sample) se zabýval Gilbert (Gilbert, Basco et al. 2002). Ukazuje, že Writing Sample má omezenou vypovídací schopnost ve vztahu k národním lékařským licenčním zkouškám. Dobře se nimi naopak korelují Biologické vědy (0,55) a též Fyzikální vědy (0,49) a samozřejmě celý test dohromady.

Zdá se, že přínos této části testu je spíš v předvídání studijního výkonu v klinických letech na lékařské fakultě (Callahan, Hojat et al. 2010), (Donnon, Paolucci et al. 2007).

MCAT(R)

V roce 1991 byl MCAT revidován. Julien se zbýval otázkou, do jaké míry MCAT skóre doplňuje informace o průměrech středoškolských známek (uGPAs) pro předpovídání úspěchu na lékařské fakultě, respektive při skládání americké lékařské licenční zkoušky (USMLE).

Ukazuje, že známky na lékařské fakultě byly dobře předpovídaný kombinací MCAT skóre a uGPAs. Při předvídání úspěchu v USMLE byly výrazně lepším prediktorem skóry MCAT a kombinací s uGPAs se predikce již jen málo zlepší (Julian 2005).

Proti zobecňování těchto závěrů (podceňování role uGPA ve srovnání s MCAT) vystoupil Kreiter ve shrnující práci (Kreiter and Kreiter 2007), který doporučuje pro předvídání akademické výkonnosti používat kombinaci MCAT a uGPA pro výběr uchazečů.

Použití testů a středoškolského prospěchu pro výběr studentů na vysoké školy

Ačkoli by se mohlo zdát, že dva široce rozšířené testy způsobí, že všechny vysoké školy budou některý z nich používat, není tomu tak. Podle údajů z roku 1999 přibližně 85% vysokých škol vyžaduje absolvování některého z testů (Schneider and Dorans 1999).

Pravděpodobně vzhledem k dlouhé historii SAT a ACT a k diskuzím, kolem které zejména používání SAT provázely, neexistuje jednotný způsob, jakým jsou výsledky testů pro přijetí používány. Nejčastěji vysoké školy kombinují výsledky SAT nebo ACT s informacemi o pořadí studenta ve třídě (HSCR) a informacemi o průměrech známek ze střední školy (GPA).

Některé státy zaručují studentům s nejlepším pořadím ve třídě (4% nejlepších) automatické přijetí na tamější státní univerzity podle jejich výběru (Lang 2007).

Záznamy o výkonu studenta ze střední školy (GPA a HSCR) jsou považovány za prokazatelné prediktory akademického výkonu (viz výše) a příspěvek SAT bývá považován za spíše diskutabilní. V některých pracích zabývajících se faktory predikujícími studijní průměry v prvním roce vysoké školy se nicméně příspěvek SAT považuje za důležitý doplněk ke středoškolským záznamům (Bridgeman, McCamley-Jenkins et al. 2000). Ukazuje se, že přínos testů jako je SAT roste, pokud je cílem vybrat excelentní studenty pro výběrové vysoké školy (BenShakhar, Kiderman et al. 1996).

Klíčový rozdíl mezi americkým systémem a například britským systémem, je v tom, že studenti v Americe absolvují zkoušky SAT nebo ACT a obdrží skóre předtím, než podají své přihlášky na vysoké školy. Jsou tedy schopni korigovat svou volbu podle výsledku testů. To je obzvláště důležité, pokud studenti uvažují o studiu na výběrové vysoké škole. Pokud je jejich skóre příliš nízké mohou změnit svůj seznam preferovaných vysokých škol, a tak získat větší šanci na přijetí. To je velmi odlišné od britského systému, kde si studenti vybírají vysokou školu na základě očekávaných výsledků státní maturity a tento odhad se může ukázat jako málo přesný.

Význam testování pomocí zkoušek SAT a ACT roste, když vezmeme v úvahu rozdílné distribuce známek na středních školách, které některým školám umožňují produkovat studenty s vyššími GPA než mají studenti z ostatních škol.

Velká Británie

Ve Velké Británii je v současnosti 32 institucí, které poskytují vysokoškolské vzdělání v oboru medicíny (UKMS 2011). Pro přijetí na školu je většinou vyžadován solidní výkon ve státních maturitách A-levels, dobrý výkon v přijímacích testech jako jsou UKCAT (vyžaduje 23 škol), BMAT (vyžaduje 5 škol) a podobně a většinou interview. Počet přijímaných studentů dosahuje ročně kolem 8000 a počet zájemců převyšuje počet přijímaných 16x (MEDICS 2005). Studenti se hlásí ke studiu ve věku kolem 18-19 let a to bez předběžného bakalářského titulu.

Pro výběr uchazečů slouží na některých školách rovněž středoškolské studijní průměry (uGPA - undergraduate grade point average), certifikát o nižším středoškolském vzdělání (GCSE - General Certificate of Secondary Education).

Standardně používaným přijímacím testem ve Velké Británii je BMAT - BioMedical Admissions Test.

Skládá se ze tří částí. První dvě části mají formu otázek s mnohočetným výběrem. Sekce 1 je zaměřena na testování schopností a dovedností (kritické myšlení, logika, uvažování). Sekce 2 je zaměřena na test odborných znalostí a schopnost jejich aplikace.

Jak diskutujeme dále v připojené publikaci, vede se o příspěvku těchto sekcí k prediktivní schopnosti testu bohatá diskuze. McManus a Emery (McManus, Ferguson et al. 2011), (Emery and Bell 2009) shodně konstatují, že příspěvky Sekce 2 a Sekce 1 v biomedicínských přijímacích testech (BMAT) jako prediktory akademického výkonu uplatňují velmi rozdílně. V citovaných publikacích autoři ukazují, že v testu BMAT používaném ve Velké Británii, Sekce 2 zaměřená na aplikace vědeckých poznatků koreluje zřetelně lépe s průměry průměru v prvním a druhém ročníku studia, než Sekce 1, zaměřená na schopnosti a dovednosti. V další práci (Emery, Bell et al. 2011) autoři ukazují, že k předpovědi studijních průměrů v prvním ročníku vysoké školy přispívá Sekce 2 signifikantně, zatímco Sekce 1 nikoli.

Vhodnost použití testu inteligence pro výběr uchazečů o studia medicíny v UK ve srovnání s vhodností použití výsledků státní maturity (A-levels) diskutoval například McManus (2005) (McManus, Powis et al. 2005). Ten shrnuje svoje předchozí poznatky (McManus, Smithers et al. 2003) i závěry jiných studií s konstatováním, že státní maturity (A level) budoucí akademickou úspěšnost predikují, ale test intelektuálních schopností nikoli.

Současně se zde konstatuje, že školy (v UK) zavádějí testy intelektuálních schopností bez důkazů o jejich vhodnosti, přesnosti či přidané hodnotě.

Široce používaným testem pro výběr studentů medicíny je UKCAT - UK Clinical Aptitude Test (test schopností a postojů).

UK Clinical Aptitude Test – je test vytvořený v roce 2006 konsorciem britských lékařských a stomatologických fakult pro testování duševních schopností uchazečů. UKCAT je navržen tak, aby testoval schopnosti a postoje, nikoli akademický úspěch, který je dobře predikován pomocí A-levels, GCSE, nebo GPA. Test je tedy zaměřen na schopnost kritického logického myšlení a schopnost vyvozovat závěry. V roce 2011 obsahoval tyto čtyři části:

- **Slovní zdůvodnění** - hodnotí schopnost uchazečů logicky uvažovat o písemné informaci a schopnost dospět k odůvodněnému závěru.
- **Kvantitativní uvažování** - hodnotí schopnost kandidátů řešit numerické úlohy.
- **Abstraktní myšlení** - hodnotí schopnost kandidátů odvodit vztahy z informací
- **Rozhodovací analýza** - hodnotí schopnost kandidátů vypořádat se s různými formami informací odvodit vztahy a přijímat adekvátní rozhodnutí.

Celá zkouška se odehrává na počítači a trvá dvě hodiny. O přínosu s efektivnosti tohoto testu se vedou diskuze (Fernando, Prescott et al. 2009).

Přínosem kognitivních a nekognitivních faktorů pro predikci klinického výkonu se opakovaně zabýval Hojat (Hojat, Borenstein et al. 1988), (Hojat, Vogel et al. 1988), (Hojat, Robeson et al. 1993), který kromě jiného na skupině 210 studentů druhého ročníku ukazuje, že přidáním psychosociálních šetření by bylo možné zlepšit prediktivní validitu MCAT vzhledem ke studijnímu výkonu na lékařské fakultě. Za poněkud nestandardní lze považovat, že psychosociální dotazníky studenti vyplňovali až ve druhém ročníku školy, vůči jejímž absolvování byla prediktivní validita měřena.

Prediktivní validitu UKCAT vůči studijnímu výkonu v prvním ročníku lékařské fakulty studoval Lynch (Lynch, MacKenzie et al. 2009). Asi aby dostal svému jménu, prokázal na velkém vzorku studentů, že UKCAT (ani jeho jednotlivé subtesty) nepředpovídají studijní výkon v prvním ročníku lékařské fakulty vůbec.

Yates na příkladu Nottingham Medical School dokázal (Yates and James 2010), že prediktivní hodnota UKCAT, zejména celkové skóre, je nízká. Test predikce celkového úspěchu bude k dispozici teprve za pár let. Pokud se však tento test mentálních schopností

nepředpovídá preklinický výkon, to je o to méně pravděpodobné, že předpovídat studijní výsledek v klinických letech.

Austrálie

Studium medicíny v Austrálii a souvisejících regionech je zpravidla 5-6 leté. Přijímání absolventů středních škol ke studiu medicíny na většinu lékařských fakult v Austrálii a na Novém Zélandu je založeno na výsledcích ze střední školy, testech Undergraduate Medicine and Health Sciences Admission Test (UMAT) a pohovoru. Před zavedením testů byli studenti vybíráni výhradně na základě studijního průměru v posledním roce střední školy. Pro přijetí k dalšímu (4 letému) studiu je vyžadováno úspěšné složení bakalářských zkoušek a dobré skóre v přijímacím testu GAMSAT.

UMAT

Test UMAT spravuje nezisková společnost Australian Council for Educational Research (ACER). Členy konsorcia je 85% lékařských fakult v Austrálii a na Novém Zélandu. Testy mají zcela svébytné složení testové sady, odlišné od dosud probíraných testů. Tvoří jej tři sekce:

1) **Test logického myšlení a schopnosti řešení problémů**

Materiály v této sekci jsou čerpány z celé řady obecných zdrojů a jsou založeny na krátkém textu, nebo informaci v grafické podobě. Otázky mají posoudit vaši schopnost pochopit a vyvodit logické závěry tím, že určíte relevantní fakta, vyhodnotit informace, najdete další nebo chybějící informace a vytváříte a testujete pravděpodobné hypotézy

2) **Test porozumění lidem**

Tato část hodnotí schopnost porozumět lidem. Otázky jsou založeny na scénáři, dialogu, nebo jiném textu představujícím konkrétní mezilidskou situaci. Otázky hodnotí schopnost identifikovat a pochopit myšlenky, pocity, chování a záměry lidí v představovaných situacích.

3) **Test neverbálního myšlení**

Otázky v této části mohou být několika druhů. Všechny jsou založeny na vzorech nebo sekvencích tvarů a jsou určeny k posouzení schopnosti uvažovat v abstraktních pojmech a řešení problémy v neverbálních kontextech.

Skóre z testu se skládá ze tří čísel určujících pořadí (percentil) za každý test. Platnost testu je dva roky. Protože test UMAT je velmi odlišný od typických zkoušek na vysokou školu nemusí vynikající akademické kvality vést nutně k vynikajícím výsledkům v testu UMAT. Test používá velká většina škol v Austrálii a na Novém Zélandu. Školy si stanovují vlastní mez pro přijetí a většina škol test doplňuje strukturovaným pohovorem a výsledky posledního roku středoškolského studia (ACER 2011).

GAMSAT

Graduate Australian Medical School Admissions Test - je test pro výběr uchazečů o magisterské studium medicíny, stomatologie a veterinárního lékařství vyvinutý v roce 1995 Australskou radou pro výzkum vzdělávání (ACER). Používá se pro výběr studentů na magisterský stupeň studia zdravotnických vysokých škol v Austrálii a od roku 1999 i na některých školách ve Velké Británii a v Irsku.

Testované znalosti jsou: Biologie a Chemie, Fyzika, Angličtina - obecné znalosti. Test trvá celý den - od osmi do šestnácti hodin. Během přestávek nesmíte opustit zkouškovou místnost.

- Oddíl I - 75 otázek z humanitních a sociálních věd (100 min.)
- Oddíl II - 2 eseje posuzující písemnou komunikaci (1 hodina)
- Oddíl III - 110 otázek z fyziky (170 min.)

Skóre se vypočítává na základě výkonnosti ve všech 3 sekcích s vahou rostoucí směrem k III. části Podle dosaženého skóre pak zdravotnické školy rozesílají pozvánky na přijímací pohovory (Coates 2008).

Izrael

Izraelský systém přijímání studentů na vysoké školy popsal Beller (Beller 1992). Izrael má na konci střední školy celostátní testy podobné státním maturitám. Na základě těchto národních zkoušek a hodnocení školy obdrží úspěšní studenti certifikát (maturitní vysvědčení - Bagrut), který je vyžadován pro vstup na vysokou školu. S rostoucím zájmem o studium na vysokých školách zavádějí některé vlastní systém výběru studentů, aby eliminovaly vliv odlišných osnov různých středních škol.

Jednotný přijímací test pro připravuje v Izraeli „Národní institut pro testování a hodnocení“ (National Institute of Testing and Evaluation). Vyvinutý test se nazývá Psychometrický vstupní test (Psychometric Entrance Test - PET). Izraelské vysoké školy definují své vlastní přijímací politiky, přesto většina z nich pracuje se skóre odvozeným z Bagrut a PET. Na některých školách mohou být navíc například pohovory nebo (při vstupu na medicínu) se vyžadují speciální oborové certifikáty.

PET

Testy PET měří kognitivní schopnosti a školní znalosti, které jsou předpokladem budoucího akademického úspěchu. Sekce testů jsou: Slovní zdůvodnění, Kvantitativní uvažování a Angličtina.

- Slovní uvažování: test sestává z otázek s mnohočetným výběrem odpovědi zaměřených na slovní zásobu, logické myšlení, verbální dovednosti a schopnost analyzovat a pochopit složité texty. Otázky mají formy slovníkového kvízu, antonym, analogií, dokončování vět. Testují logiku, deduktivní a induktivní uvažování a čtení s porozuměním.

Studie Tamar Kenet-Cohen (Kennet-Cohen, Bronner et al. 1999) zkoumala prediktivní validitu jednotlivých částí testu PET na šesti vysokých školách v Izraeli a zjistila, že verbální test je lepším prediktorem než další dvě části testu. Při podrobnějším pohledu se ukazuje, že ke kvalitě verbálního testu nejméně přispívají „analogie“, lépe jsou na tom „čtení s porozuměním“ a test „induktivně deduktivního myšlení“. Obecně přispívaly k predikci studijního výkonu nejtěžší otázky nejvíce. Naopak nejsnazší otázky týkající se slovní zásoby, dokončování vět a podobně měly k predikci až negativní příspěvek.

- Kvantitativní uvažování: test sleduje schopnost uchazečů používat čísla a matematické koncepty. Typy úloh jsou algebraické problémy, rovnice a geometrické problémy. Vyžadovány jsou pouze základní znalosti z matematiky, pokud jsou potřeba složitější vzorce, jsou uvedeny v brožuře přiložené k testům.

- Angličtina: test má posoudit znalost anglického jazyka v míře považované za nezbytnou pro čtení akademické texty. Otázky jsou zaměřeny na dokončování vět, čtení a porozumění.

Výsledkem testů je skóre s průměrem 500, směrodatné odchylkami 100 a mezemi od 200 až do 800 bodů. PET velmi koreluje se SAT (0,82), a to zejména v matematické části (0,85), ale korelace mezi PET a Bagrut jsou mnohem menší, obvykle kolem 0,45. V současné době je PET přeložený do arabštiny, ruštiny, angličtiny, francouzštiny a

španělštiny, s kromě anglické části testu. Jones (Jones 1994) poznamenal, že PET sdílí mnoho společných rysů se SAT, zvláště z pohledu struktury. Pomocí faktorové analýzy bylo zjištěno, že PET (podobně jako SAT) hodnotí dva faktory, slovní a matematické uvažování. Tyto faktory odpovídají „fluidní inteligenci“ (schopnost flexibilního řešení problémů) a „krystalické inteligenci“ (schopnost využít nahromaděné znalosti) jak je užívána k popisu schopností uchazečů jinde v literatuře.

Skóre z testu PET předpovídá výkon v prvním ročníku lépe, než známky ze střední školy. Jones (1994) argumentoval, že to může být souviset s tím, že test PET je časově blíže vstupu na vysokou školu.

Celý test má v průměru 252 otázek, z nichž 168 slouží pro potřeby hodnocení. To odpovídá tomu, že osmi částí testu jen šest má vliv na hodnocení. Zbývá třetina otázek (dva bloky testu) se označují jako „pilotní“ a slouží ke kontrole kvality nových otázek a pro potřeby porovnání obtížnosti zkoušek v různých termínech.

Problémy PET

Téměř polovině arabských studentů, kteří složili maturitní zkoušky (Bagrut) se nepodařilo dostat na vysokou školu z důvodu špatných výsledků v psychometrických testech, což je zarážející ve srovnání s pouze 20% neúspěšných židovských žadatelů. Rozdíl v psychometrických hodnoceních židovských a arabských studentů zůstává stabilní je více než 100 bodů z celkových 800. Testy jsou proto napadány jako kulturně předpojaté, nicméně, Zeidner ve své studii (Zeidner 1986) zjistil jen nepatrné rozdíly v prediktivní validitě testu napříč různými kulturními skupinami, a tak se zdá, že pozorovaný rozdíl souvisí spíše s psychometrickými než kulturní rozdíly.

Shrnutí literární rešerše

Zabývali jsme se metodami výběru uchazečů o vysokoškolské studium se zvláštním přihlédnutím k výběru studentů medicíny. Zahraniční zkušenosti můžeme shrnout do několika bodů:

1. Pro výběr budoucích studentů je obvykle využívána kombinace všech dostupných informací, zvláště pak studijních výsledků ze střední školy a výsledků přijímacích testů.
2. Testy zadávané v různých termínech bývají mezi sebou provázány pomocí otázek, které se porovnávaných testech opakují a neslouží primárně k hodnocení studentů, nýbrž k hodnocení otázek samotných a porovnání obtížnosti testů.
3. Prediktivní validita přijímacích testů je ve shromážděné literatuře téměř výhradně vztahována ke studijním průměrům v prvních ročnících vysoké školy (nikoli tedy k celkovému studijnímu úspěchu).
4. Prediktivní validita schopnostních (a psychologických) testů je trvale diskutována, zatímco prediktivní validita znalostních testů je nesporná.

Předběžná studie

Vlastní příspěvek ke studované problematice je v této práci zpracován formou příložených publikací a samostatných analýz. Nejprve proto stručně shrneme výsledky předběžné studie zabývající se možnostmi využít znalost studijních výsledků na střední škole pro výběr uchazečů o studium medicíny. Motivací k této studii byl požadavek vedení fakulty ověřit potenciální přínos přijetí části studentů na základě jejich výsledků ze střední školy.

Studium prediktorů úspěšného absolvování studia na 1. LF UK

Cíle

Cílem naší studie bylo zhodnocení vlivu jednotlivých faktorů známých v době přijetí na vysokou školu na úspěšnost studia medicíny. Úspěšností zde nazýváme schopnost úspěšně ukončit studium v řádném termínu, tedy do šesti let od začátku studia.

Datový soubor

Předběžná studie byla provedena na datech studentů přijatých v akademickém roce 1996/97 (kdy byly naposledy u části studentů k dispozici všechny známky ze střední školy). Základní soubor tvoří 267 studentů přijatých 1996, u kterých byly dostupné všechny údaje o studiu a známkách na střední škole. Z toho 94 ukončilo fakultu úspěšně, 117 ukončilo neúspěšně, 56 k datu sběru dat ještě studovalo (a nebyli do dalších výpočtů zahrnuti). U sledovaných studentů známe výsledky testů F, B, Ch a ručně jsme sesbírali známky z M, F, Čj, Bi, Ch. Poznamenejme, že podklady pro tento sběr dat byly vzhledem ke stavu archivu značně neúplné a procento studentů, kteří k datu sběru ještě studovali, bylo nezanedbatelné. Skupinu studentů, kteří školu k danému datu dokončili, značíme Fin-u, skupinu studentů, kteří byli ve studiu neúspěšní značíme Fin-n. Soubor obsahoval proměnné:

ST1,...,ST4	průměr všech známek v prvním až čtvrtém ročníku
ST1_4	průměr z předchozích čtyř průměrů
M1...M4, F1...F4, CH1...CH4, B1...B4, CJ1...CJ4	známky z matematiky, fyziky, chemie, biologie a českého jazyka po jednotlivých ročnících
M1_4, F1_4, CH1_4, B1_4, CJ1_4	průměr z předchozích čísel za čtyři roky
F, B, CH	tři části přijímacích zkoušek, body za chemii, biologii a fyziku
BODY	součet F+B+CH
STAV	informace, zda student úspěšně skončil školu v řádném termínu
SKOLA	informace o typu střední školy studenta

Zajímala nás možnost predikce výsledku studia na základě těchto dostupných dat.

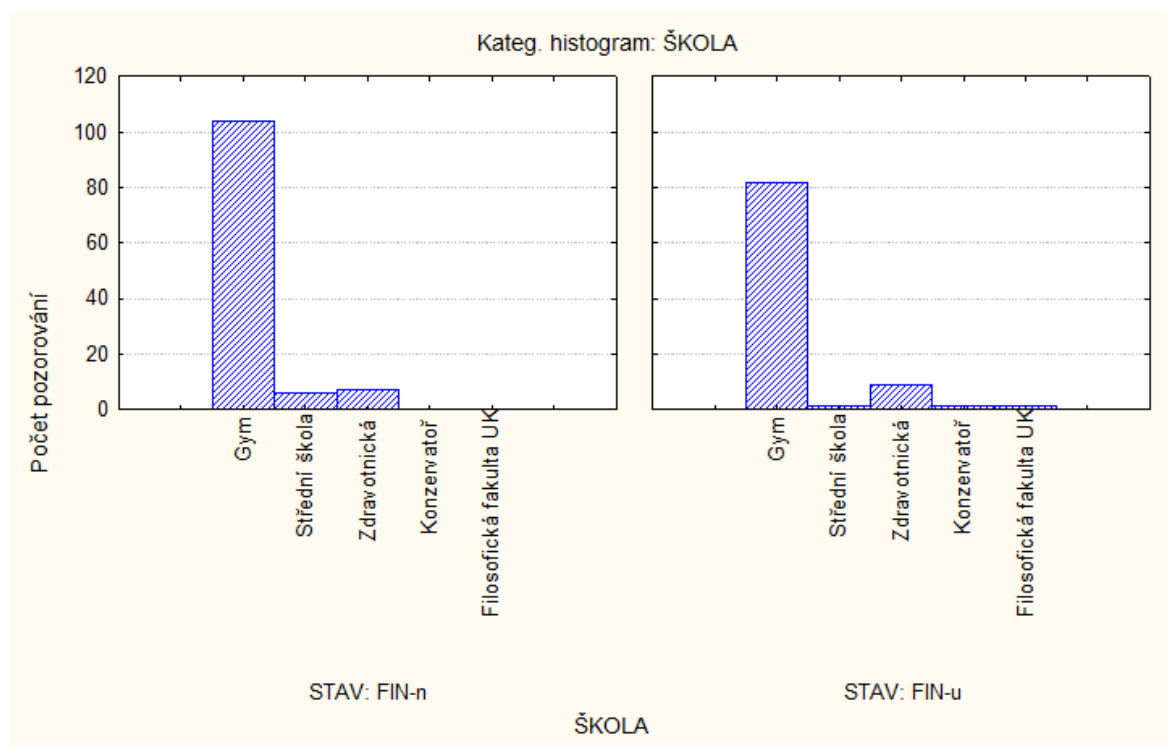
Popisné charakteristiky

Pro sledované proměnné jsme zpracovali popisné statistiky.

Tabulka 3. Popisné charakteristiky sledovaného souboru studentů.

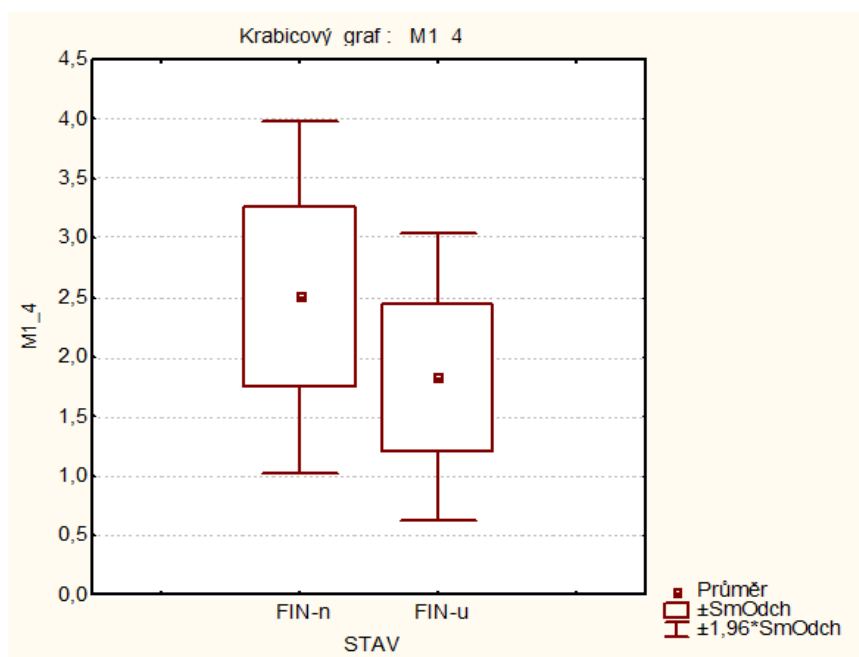
	N	Minimum	Maximum	Mean	Std. Deviation
ST1	210	1,00	3,00	1,7010	,46151
ST2	211	1,00	3,30	1,6888	,48768
ST3	211	1,00	3,50	1,6611	,48275
ST4	211	1,00	3,50	1,6540	,48674
M1	203	1	4	2,16	,847
M2	204	1	4	2,28	,870
M3	196	1	4	2,21	,878
M4	192	1	4	2,17	,866
M1_4	204	1	4	2,21	,773
F1	204	1	4	2,11	,808
F2	205	1	4	2,00	,869
F3	189	1	4	1,97	,821
F4	162	1	4	1,86	,779
F1_4	209	-4,0	-1,0	-1,985	,6928
CH1	203	1	4	1,97	,838
CH2	202	1	4	1,89	,809
CH3	188	1	4	1,64	,758
CH4	94	1	4	1,63	,733
CH1_4	208	1,0	4,0	1,825	,6775
B1	168	1	4	1,63	,672
B2	192	1	4	1,62	,691
B3	188	1	4	1,48	,658
B4	163	1	4	1,56	,721
B1_4	205	1,0	3,7	1,583	,5639
CJ1	209	1	4	1,85	,737
CJ2	211	1	4	1,88	,746
CJ3	211	1	4	1,89	,751
CJ4	209	1	4	1,89	,836
CJ1_4	211	1,0	4,0	1,893	,6620
CH	211	30	97	74,10	13,162
B	211	39	100	85,43	10,624
F	211	43	100	81,28	11,978
BODY	211	139	296	240,82	31,661
ST1_4	210	1,00	3,21	1,6779	,45147
Valid N (listwise)	81				

Na kategorizovaném histogramu budoucích úspěšných a neúspěšných studentů podle typu absolvované střídání školy můžeme vidět, že převažujícím typem střední školy v obou skupinách bylo gymnázium. Rovněž v obou skupinách byli zastoupeni rovněž absolventi středních zdravotnických škol, ale nikoli již absolventi ostatních středních škol, kteří ukončili studium lékařské fakulty vesměs neúspěšně.

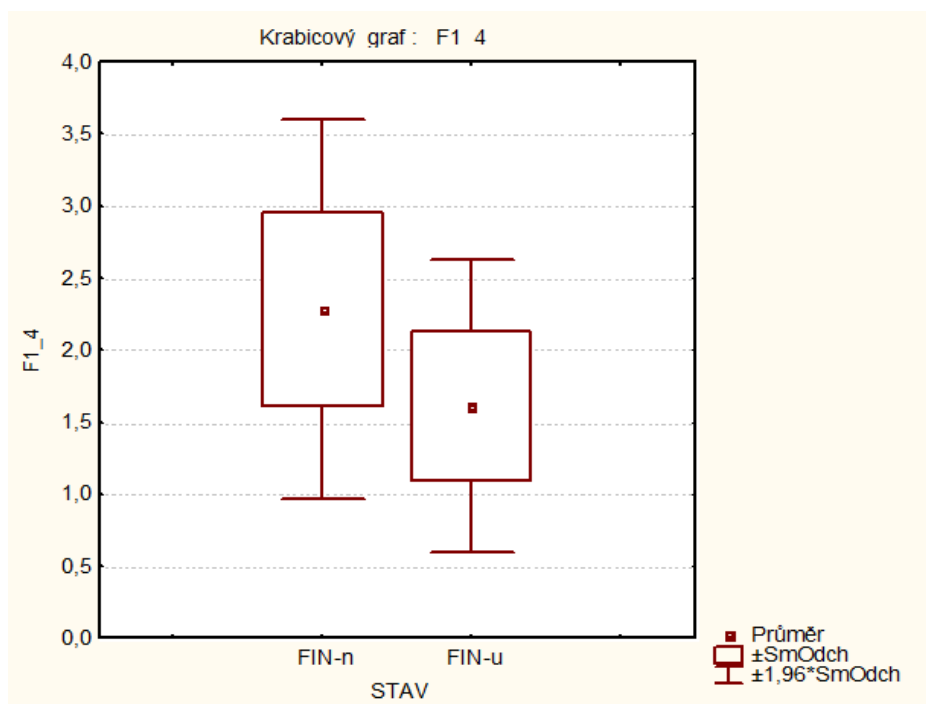


Obr. 5. Kategorický histogram podle typu střední školy.

O vypovídací schopnosti známek ze střední školy svědčí např. grafické znázornění průměrů středoškolských známek pro budoucí úspěšné a neúspěšné studenty medicíny pomocí krabicových grafů.

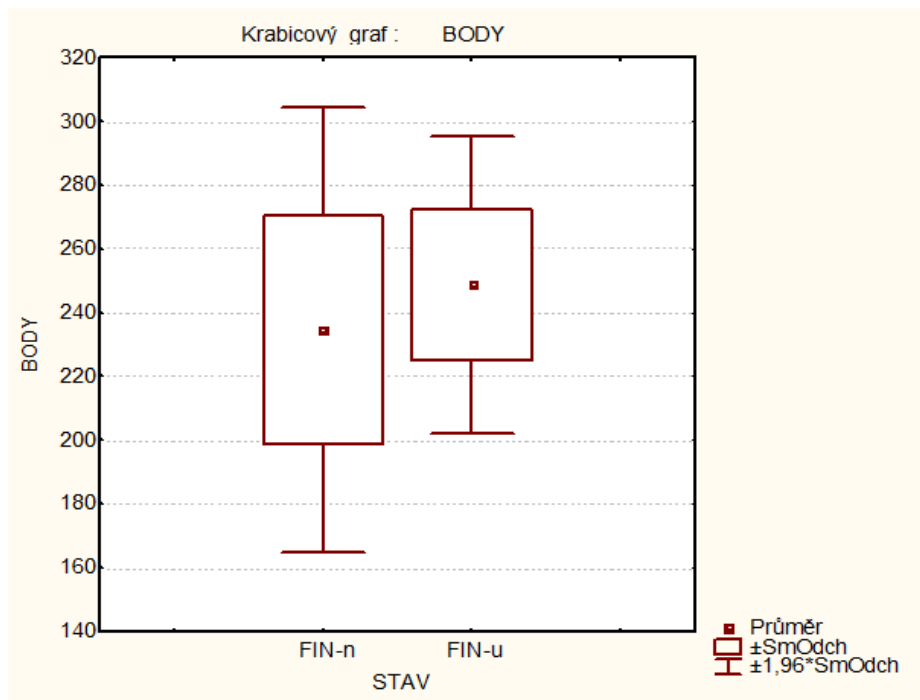


Obr. 6. Krabicový graf průměrů známek z matematiky za celé středoškolské studium.



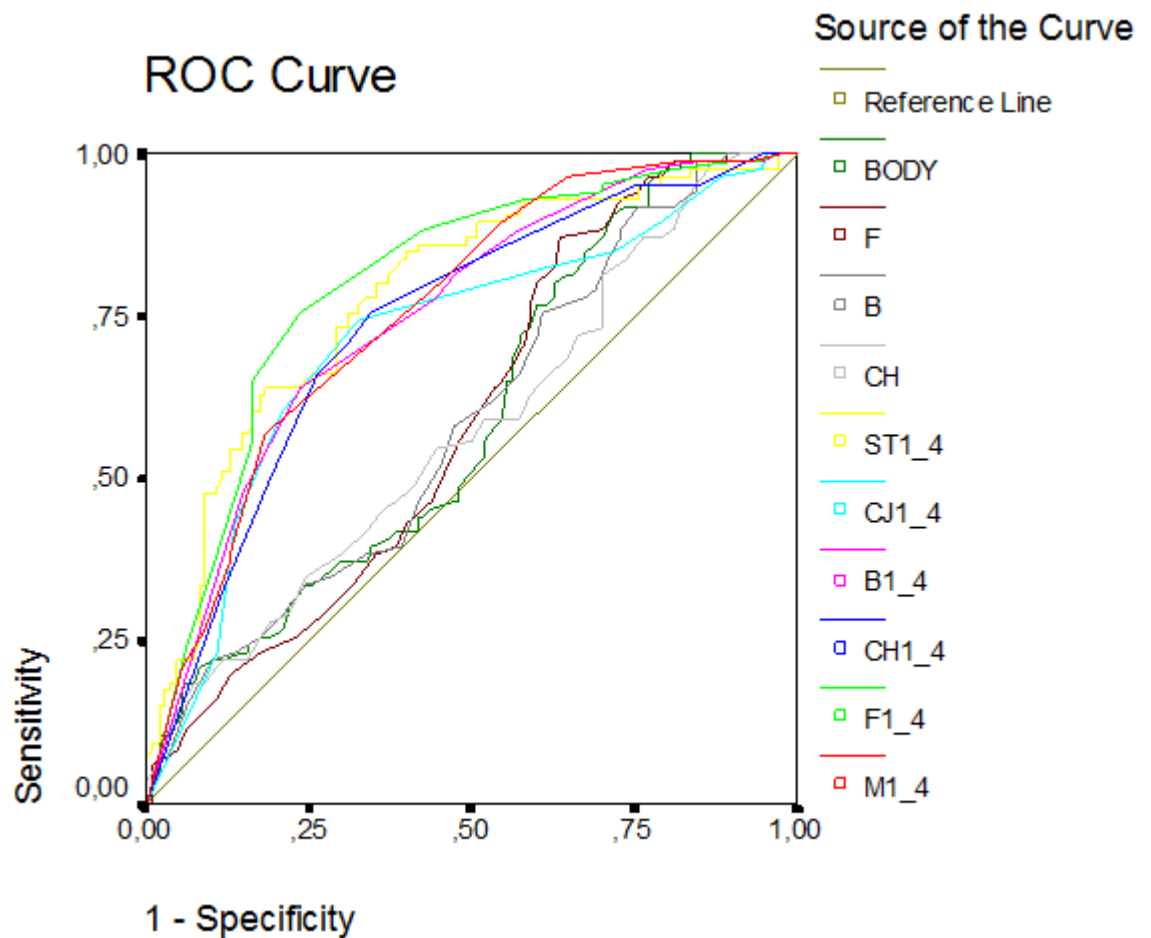
Obr. 7. Krabicový graf průměrů známek z fyziky za celé středoškolské studium.

Naproti tomu krabicový graf budoucích úspěšných a neúspěšných studentů rozlišených podle počtu dosažených bodů v přijímacích testech ukazuje, že test stratifikuje spíše ve své spodní limitě.



Obr. 8. Krabicový graf porovnání průměrů bodových hodnocení z přijímacích testů pro obě skupiny.

Pro vyhledání nejlepší vhodných prediktorů jsme použili metodu ROC křivek.



Obr. 9. Porovnání ROC křivek v závislosti na jednotlivých prediktorech.

Pro ROC křivky významných prediktorů jsme vypočetli plochy pod křivkou:

Tabulka 4. Porovnání prediktorů pomocí metody ROC křivek.

prediktor	M1_4	F1_4	CH1_4	B1_4	CJ1_4	ST1_4	CH	B
plocha pod křivkou	0,759	0,8	0,738	0,751	0,715	0,782	0,577	0,587

Z uvedeného porovnání se jeví, že známky ze střední školy (např. průměry známek za celé studium) patří k dobrým prediktorům budoucí úspěšnosti studia.

Závěr

Uvedená studie přinesla předběžné výsledky o vhodnosti využití středoškolských studijních průměrů pro predikci budoucího úspěchu při studiu medicíny. Vzhledem k nedostupnosti úplných vstupních dat byl však její přínos spíše v ověření metodiky.

Podrobněji se proto tomuto tématu věnujeme v následující části, kde formou publikace analyzujeme data studentů přijatých v roce 1999 ke studiu na 1. LF UK v Praze.

U těchto studentů již sice nebyly k dispozici známky z jednotlivých předmětů středoškolského studia, jen studijní průměry za celé studium, ale podle naší předběžné studie právě tyto studijní průměry patří k nadějným prediktorům.

The prediction and probability for successful completion in medical study based on tests and pre-admission grades

Authors: Čestmír Štuka¹, Patrícia Martinková², Karel Zvára², Jana Zvárová²

Institutions:

¹ Computer Technology Centre, First Faculty of Medicine, Charles University in Prague, Czech Republic

² Centre of Biomedical Informatics, Department of Medical Informatics, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague

ABSTRACT

The study analyses admission procedure at the largest school of medicine in the Czech Republic. Based on 1999 data of 383 students admitted that year, it should be noted that 163 of them failed and 220 have succeeded to finish the study. By use of Pearson's correlations between predictors, cluster analysis and logistic regression we show that preadmission grades predict the overall success in medical study with the same accuracy as admission tests but each of them describe different dimension of students' abilities. Simultaneous use of GPA and AT in admission process of all students should bring higher quality of selection process and also practical advantages for future analyses. Nevertheless, the idea of having a group of the students admitted based only on an excellent GPA, could be considered a quite reasonable one

Acknowledgment: The paper was supported by project 1M06014 by MŠMT ČR and by the institutional research plan AV0Z10300504. Authors would like to thank to all who contributed to the manuscript with their valuable advice, comments and editorial input.

Key words: education, admission criteria, pre-admission grades, admission test, medical study

INTRODUCTION

The proper selection of students to the higher education is crucial point both for the applicants and the institution: the quality of students influences the school's reputation and vice versa. The number of applicants often exceeds available spaces and admission committees face a very difficult and challenging task. Developing a clear definition of suitability for medical training is the first priority, whether locally or nationally (Parry et al, 2009).

There have been many studies written on predictors of medical academic performance. Salvatori (2001) made a review of 83 articles from field of medical education and examined various selection tools. She concluded that the pre-admission overall grade-point average (GPA) is clearly the best predictor of academic performance.

Medical College Admission Test (MCAT) - standardized test given in USA, Canada and 15 other countries, was found to be a good predictor and Salvatori cited studies, where MCAT scores together with GPA accounted for 21% of the variability in academic in-course grades. Although GPA is single best predictor of academic achievement, much of variability of academic performance still remains unexplained. However, the evidence to support the continued use of personal interviews, essays or letters of reference remains controversial; further research is needed to find more reliable and valid ways of accessing the non-cognitive characteristics of applicants.

The relationship of GPA to clinical performance is even less clear (Salvatori, 2001). Other, perhaps non-cognitive variables are contributing to clinical performance. Firstly, students have already learned how to study at university. Secondly, in clinical subjects the requirements for memorizing are not so high, while the accent is put on logical thinking and interconnecting knowledge from different subjects. Students also newly come into personal contact with patients and their motivation and level of responsibility changes.

Emery and Bell (2009) studied the predictive validity of the BioMedical Admissions Test (BMAT) used in the United Kingdom. As a reaction, McMannus et al. argued that without evidence of incremental validity it is difficult to assess the value of any selection tests for medicine (McMannus et al. 2011a). Mentioned publications describe attitudes in countries with different cultural background and different education system, thus their results cannot be transferred without further verifications nevertheless the requirements on incremental validity of tests can be generalized.

Entry to Czech medical schools is very competitive task. For example, at First Faculty of Medicine of the Charles University in Prague (1st LF UK) in the period of 2006 – 2011 the number of applications to study of general medicine exceed in average 2.5 times number of admitted students (derived from Línová, 2011). We should consider the fact that students often apply for more schools and if they are admitted they have the possibility to choose. Nevertheless, schools do count with this “inflation” (at 1st LF UK, in the given period, only 53% ± 0,06% of admitted students started the study) and they admit more students than the spaces they have. In the result, more than one half of the applicant for the study of general medicine at 1st LF UK remains unsatisfied.

In the Czech Republic, conditions for acceptance for higher medical education differ greatly between universities, but also between university faculties and sometimes even between study branches inside the same faculty. Traditionally, faculties organize in-house admissions consisting mainly of knowledge tests, general aptitude tests and/or interviews. Some faculties take into account students’ pre-admission performance (mostly measured by GPA). While in the United States, the Great Britain and other countries, examination of relationship between admission test performance and subsequent study performance is part of the test evaluation, in the Czech Republic, this issue has rather been neglected; the changes in admission process often take place without previous analysis of quality of existing tests (Rubešová, 2007).

Admission process to medical schools in the Czech Republic is often discussed at domestic conferences (Štuka and Smutek, 2003), but there were and are only few works finally published in journals (national or even international). Admissions to Czech medical schools were systematically studied by composite authors Höschl, Kožený and Tišanská who analysed admission process to the Third Faculty of Medicine of the Charles University in Prague. Study at 3rd LF UK is compatible with other medical faculties, but unlike other faculties it is in done by PBL (problem based learning). Therefore the conclusions made on study performance predictors at 3rd LF UK should be applied at other schools of medicine only with caution.

In this article, we analyse admission process to 1st LF UK and similarly to the work of Höschl and Kožený (1997) and work Štuka and Šimeček (2006), we try to identify variables predicting overall academic success and to find the optimal admission criteria for medical study. Compared to mentioned work and work of Kožený and Tišanská (2001) we

do not look for the high-school subject, which grades best predict success in medical study, but rather investigate whether subject GPAs give more information than overall GPA. Beside of searching for relationships between different predictors and estimation of their predictive power, we centralize on verification of the current two-step admission and on relevance of the criteria for offering place at 1st LF UK without admission tests. The aim of the study was:

1. To assess the contribution predictors for successful study bring, based on total secondary/high school grades; contribution of predictors based on individual parts of medical college admission test, and to compare the validity of their predictability in both instances.
2. With the help of such predictors to find an optimal model predicting success in medical studies.
3. To assess and evaluate criteria used during students' admission based on previous academic performance and high school grades.

METHODS

Subjects

The study is based on data of students who were admitted in 1999 to 6-year Master course of General Medicine held in Czech language at 1st LF UK. Year 1999 was the last year when all the students must have passed admission tests to get admitted to 1st LF UK. In total, 467 students started the freshmen year (386 Czech, 75 Slovak students and 6 other foreigners with average age of 19.88 (\pm 2.18)). Those students were excluded from analysis who were admitted to 1st LF UK repeatedly (24 students), whose data about year of high-school graduation (10) and overall GPAs (17) were not complete and who did not possess grades from high school (56 students). In total, 383 students were analysed. Out of them were 316 Czech and 67 Slovak, 253 women and 130 men, 348 studied Gymnasium and 35 different type of high school.

Admission conditions

At 1st LF UK, the admission process changed during past decade. While earlier, it has been based on total score of admission test (AT) consisting of test in biology (ATb),

chemistry (ATc) and physics (ATp), since 2002, good pre-admission study performance may give a student an advantage of being accepted without AT. Approximately 45% of available space is offered to students with best pre-admission GPA without requirement on passing AT. Beside excellent average high-school GPA (better than 95 on scale 100=excellent to 0=insufficient), three additional criteria must be fulfilled by students, who want to apply for acceptance without AT:

A1 Successful graduation from high-school in the year of application (elimination of repeated attempts)

A2 Completion of two semesters of course in profile subjects (biology, chemistry, physics) within the last two years of high-school study with final grade not less than 75 (students and type of high school on science)

A3 Passing the Czech language graduation exam (program is held in the Czech language)

Maximal number of students admitted without AT is determined and made public in advance (for year 2011 it was 210 out of approx. 470 admissions). Applicants with best overall GPA satisfying conditions A1 - A3 are given an offer for 1st LF UK. Other applicants are invited to admission test.

Since 2009, General Aptitude Test has been added into admission process for three years, but it is not subject of this work and a separate publication is planned on this topic.

Measures in the data

High school performance

For all applicants, **type of high school** and **year of graduation** (corresponding to condition **A1**) were recorded. Final grades in math, physics, biology, chemistry and the national language courses were recorded for all four years of high-school study. Out of them, mean scores over the 4 years were enumerated and assigned as **GPA_m**, **GPA_p**, **GPA_b**, **GPA_c**, **GPA_{cz}**. Also, accomplishment of condition **A2** was checked. Overall **GPA** based on all courses taken by student within the 4 years of study was recorded.

Written entrance examination (admission test, AT)

The admission test consisting of chemistry, physics and biology subtest, each included 100 multiple-choice items. One point was assigned for each item in case of choosing all (and only) correct answers; giving maximum of 100 points for each sub-test (assigned **AT_c**, **AT_p**, **AT_b**) and 300 points for the whole test (**AT**). Items for each sub-test were chosen randomly from set of approximately 1000 items, most of which were published either in exact wording or its variation. All of the students taking the test in the same day took the same test, but the order of items in the test and of given answers were different.

Other information about student

For all applicants, beside other information, their **sex** and **nationality** were recorded. Czech nationality was used as an indicator of fulfilment of condition A3 (Czech language graduation exam is mandatory for all Czech high school students).

Success in the medical study

The predicted variable was student's overall success in the medical study. Usual length of medical study at 1st LF UK is 6 years (suggested as predicted variable in Höschl, Kožený (1997) and used in Kožený, Tišanská (2001)), but recently many students take advantage of exchange programs and extend their study, we had therefore waited until the study result (finished successfully/failed to finish) was known for all students .

Statistical analysis

The strength of relationship between predictors was estimated by *Pearson correlations*. The dissimilarities between predictors were analysed by *cluster analysis*; Euclidean distance was used as metrics. *Stepwise logistic regression* based on Akaike's information criterion (*AIC*) was used to find the best logistic model for prediction of success in medical study. *Likelihood-ratio test* was used for testing equivalence between chosen models of logistic regression. *Contingency tables* of true success in medical study and success predicted by different combination of variables were computed and compared by means of *odds ratio (OR)*.

All statistical analyses were performed with the help of the statistical environment R (R 2010) and its libraries *rms* (Harrel, 2012), *MASS* (Venables & Ripley, 2002) and *Epi* (Carstensen et al., 2011).

RESULTS

Out of 383 analysed students admitted in year 1999 to 1st LF UK, 220 succeeded to graduate and 163 failed. Detailed description of students by their academic success is included in Table 1 and Table 2.

Table 1: Characteristics of students by their academic success at 1st LF UK

	Succeeded (N=185)		Failed (N=198)	
Sex	145 Female	75 Male	108 Female	55 Male
Type of high school	212 Gymnasium	8 Other	136 Gymnasium	27 Other
Year of HS graduation (A1)	174 True	46 False	78 True	85 False
profile B+C+P (A2)	159 True	54 False	57 True	106 False
Czech graduation exam (A3)	172 True	48 False	144 True	19 False

Table 2: Descriptive statistics of GPA and AT predictors by academic success at 1st LF UK

	Succeeded (N=220)					Failed (N=163)				
	mean	SD	Min	max	N	mean	SD	min	Max	N
GPAm	73.75	19.32	25.00	100.00	219	65.19	19.41	25.00	100.00	161
GPAp	80.95	16.83	25.00	100.00	217	69.67	18.69	25.00	100.00	161
GPAc	83.58	14.76	37.50	100.00	217	73.62	17.89	33.33	100.00	160
GPAb	91.25	10.93	58.33	100.00	218	82.51	15.06	43.75	100.00	154
GPAcz	82.22	14.17	43.75	100.00	220	74.22	17.51	31.25	100.00	163
GPA	87.14	9.43	62.50	100.00	220	79.26	12.71	47.50	100.00	163
ATc	72.54	15.33	34.00	99.00	220	64.28	14.37	27.00	97.00	163
ATb	82.08	11.56	51.00	99.00	220	76.12	11.04	51.00	99.00	163
ATp	78.33	13.74	42.00	100.00	220	69.77	13.82	35.00	100.00	163

AT	232.94	37.00	166.00	298.00	220	210.18	32.87	166.00	295.00	163
-----------	--------	-------	--------	--------	-----	--------	-------	--------	--------	-----

min minimum value, **max** maximum value, **SD** standard deviation, **GPA** high-school grade-point average, **m** mathematics, **p** physics, **c** chemistry, **b** biology, **cz** Czech/Slovak language, **AT** admission test

Relationship between different GPA and AT predictors

Pearson correlations between different GPA and ATs are displayed in Table 4. We can see that ATs give quite different information about student than GPAs (low values of correlation coefficient in upper right part). On the other hand, information given by different ATs is quite similar to each other (high values in lower right part), and information given by single GPAs is similar to each other, too (high values in upper left part).

Table 3: Pearson correlations between predictors

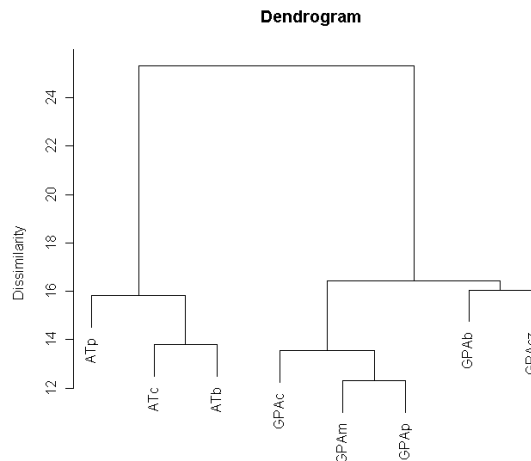
	GPAm	GPAp	GPAc	GPAb	GPAcz	GPA	ATc	ATb	ATp	AT
GPAm	1.00	0.79	0.75	0.64	0.70	0.84	0.24	0.19	0.19	0.23
GPAp		1.00	0.77	0.71	0.63	0.83	0.24	0.20	0.20	0.24
GPAc			1.00	0.71	0.68	0.82	0.24	0.22	0.17	0.24
GPAb				1.00	0.65	0.79	0.20	0.20	0.13	0.20
GPAcz					1.00	0.83	0.20	0.20	0.14	0.20
GPA						1.00	0.26	0.24	0.22	0.27
ATc							1.00	0.74	0.67	0.91
ATb								1.00	0.66	0.88
ATp									1.00	0.88
AT										1.00

GPA grade-point average, **m** mathematics, **p** physics, **c** chemistry, **b** biology, **cz** Czech/Slovak language, **AT** admission test

Dissimilarities between predictors are displayed in dendrogram of cluster analysis (Figure 1). Measure of dissimilarity is the height at which they join a single group. Again, we can see that ATs give different information about students' quality than information given by GPAs. Looking more closely at dendrogram, beside big difference between ATs and GPAs (two parts of dendrogram), inside ATs we can see a dissimilarity between

physics and other two AT subtests (biology and chemistry) which are not concentrated so much on logical thinking but rather on memorizing. Similarly, GPAs in physics and mathematics describing student's logical thinking are the two most similar, connecting further with chemistry and standing apart from biology and Czech/Slovak language.

Figure 1: Cluster analysis dendrogram.



Optimal model for predicting study success

Having only scores from three admission tests ATc, ATp and ATb, we might look for their optimal weights for prediction of academic success at 1st LF UK. Such model accounts for 13% of the variability in academic success (Nagelkerke $R^2=0.128$). Nevertheless, we cannot reject the submodel of prediction of academic success by their sum AT ($p=0.498$). Such submodel accounts for only slightly less variability (Nagelkerke $R^2=0.124$). Thus, equal weighting of subtests is reasonable.

Another question might arise - whether all three subtests are needed. Analysis show, that important information is carried only by physics subtest which by itself explains 12% of variability (Nagelkerke $R^2=0.115$). Both other subtests – chemistry and biology may be omitted without significant information loose ($p=0.126$).

Similar predictability (and even higher than AT) possesses also high-school performance: total GPA accounts for 15% of the variability in success at 1st LF UK. The same percentage of variability is explained by optimal combination of performance in single subjects GPAc, GPAm, GPAP, GPAb and GPACz.

As was shown earlier, AT and GPA describe different parts of student abilities and thus their combination should have better prediction ability than AT and GPA themselves. Logistic regression for prediction of study success by combination of AT and GPA gives model

$$\text{Logit}(P(\text{success})) = -7.745 + 0.015 \text{ AT} + 0.056 \text{ GPA} \quad (\text{model 1})$$

According to this model, optimal weight of AT (with max. 300 points) to GPA (max 100 points) is 45% to 55%. This model accounts for 22% of the variability of success at 1st LF UK.

Let us now try to find optimal predictor of student's success in medical study based on all accessible variables. All five categorical predictors mentioned in Table 1 and all ten continuous predictors displayed in Table 2 were taken into the null model of logistic regression for predicting probability of success in medical study. Automated stepwise model selection was processed: in each step, one predictor was removed with aim to maximize Akaike's information criterion (AIC), criterion of optimal prediction which penalizes number of predictors. Resulting model contains these variables: type of high school, criterion A1 (year of graduation), criterion A2 (profile classes), admission test score (AT), performance at high school represented by GPA and GPA_m. After removing insignificant predictors – criterion on profile classes (highly dependent on type of the high school and on GPA) and GPA_m (strongly correlated with total GPA) we get optimal model for prediction of study success at 1st LF UK:

$$\text{Logit}(P(\text{success})) = a_0 + a_1 \cdot (\text{year} = 1999) + a_2 \cdot (\text{type of HS} = G) + a_3 \cdot \text{AT} + a_4 \cdot \text{GPA} \quad (\text{model 2})$$

with estimates given in Table 4.

Table 4 Coefficients of optimal model of prediction of success at 1st LF UK

Coefficient	Estimate of a_i	SE(a_i)	Z	p value
a_0	-8.973	1.255	-7.147	<0.001
a_1 (year = 1999)	1.082	0.279	3.880	<0.001
a_2 (HS type = G)	1.164	0.465	2.502	0.012
a_3 (AT)	0.018	0.004	5.082	<0.001
a_4 (GPA)	0.042	0.012	3.638	<0.001

Model accounts for 31% of the variability in success. According to this model, optimal weight of AT (with max. 300 points) to GPA (max. 100 points), criteria on year of graduation (max. 1 point for year 1999) and type of high school (max. 1 point) is 46% to 35% to 10% to 9%. Compared to model 1, the impact of conditions on year of graduation and type of high school is at the expense of lowered impact of GPA.

Current admission criteria

Current admission criteria to 1st LF UK were set up in 2001 more or less on the basis of estimation and experiences of admission committee. Nevertheless, they greatly correspond with results given in previous part of our analyses, where GPA, AT and conditions A1 and A2 (dependent on GPA and type of high school) were selected as the important ones for prediction of success.

Nevertheless acceptance to 1st LF UK is not decided by weighted sum of GPA, AT, A1 and A2. Instead, the best students according to GPA and satisfying conditions A1, A2 and A3 are given an offer to 1st LF UK without passing AT. And the AT score decides only on admission of the rest of students. This saves costs for admission process and encourages the students with best GPA to choose 1st LF UK.

New admission criteria led to quite different population of students from population admitted in 1999 and earlier: students who might have been admitted today without passing AT, may have not been admitted in 1999 - because of not passing AT, or even not trying to pass AT.

In recent years, approximately 210 out of 450 available spaces are offered to students with best GPA without requirement on passing AT. In year 1999, only 77 out of 383 analyzed students (20%) had $GPA \geq 95$, and only 32 students (8%) satisfied also the three additional criteria and would be given an offer for 1st LF UK without AT today. As a consequence, average GPA of students admitted in 1999 was probably somewhat worse than the current one. Bearing this limitation in mind, in following we try to show, how the current criteria for admission without AT improve the study success at 1st LF UK.

In Fig. 2 we compare the current criterion for admission without AT (A1+A2+A3, black line) with criteria not containing A1, A2, or A3. The higher the odds ratio OR is, the more suitable is the criterion. For less strict GPA requirements, eliminating A3 (red line) seems to be profitable, nevertheless for more strict GPA requirements (1st LF UK requires $GPA \geq 97$) eliminating of condition A3 causes lowering of the odds ratio and thus the current requirements A1+A2+A3 may be proved as reasonable.

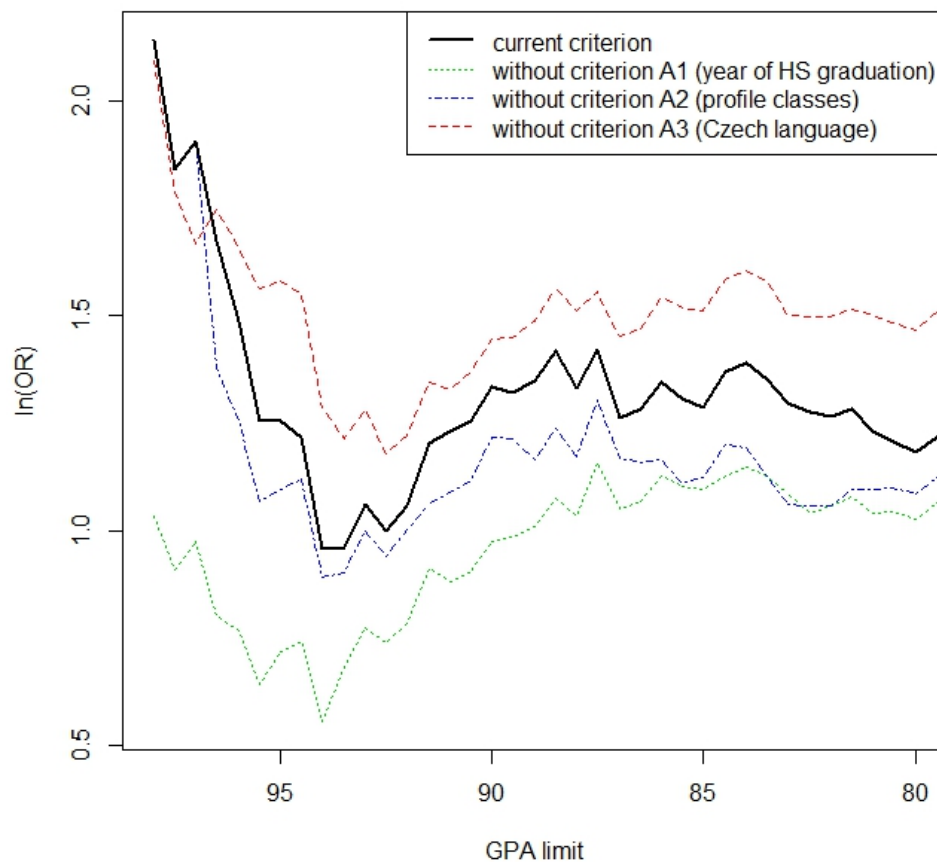


Figure 2: Comparison of criteria combinations involving different GPA limits on the basis of logarithm of odds ratio.

DISCUSSION

In this study, preadmission grades (GPA) were shown to predict overall success in medical study almost with the same accuracy (and even a bit higher) as admission tests. This is true despite of existing differences between high schools and their quality/difficulty.

Admission tests and pre-admission grades were shown to describe different dimensions of students' abilities. While preadmission GPA describes long-term study success and ability to succeed in wide scale of high-school subjects, as the opposite, the admission tests cover only knowledge in profile subjects and are cross-sectional (measured at the moment) and thus might be affected by many biases (caused by student's actual health condition, examination fever, nervousness etc.). On the other hand, they describe student's ability to learn large amount of information in limited time, which might be crucial for future medical studies. And unlike GPA, which might be affected by school's quality/difficulty, the admission tests are equal for all students, or their equal difficulty can be easily ensured in the case of more versions of the test.

Beside predictive validity of admission tests for 1st LF UK, also their incremental validity was proved by this study, as required by McManus et al. (2011a) for psychometric data reports. We have proved that admission tests give new information about students' predispositions to succeed in medical study **over** information given by preadmission GPA. Question remains, whether the incremental validity of current admission tests will be proved also in future with respect to central leaving examinations, newly introduced in the Czech Republic. There are doubts however, whether our current GCSE equivalent exam in its present form will remain usable, with respect to aggregate subjects tested for during medical college admission tests.

Also, it was shown that overall GPA describes ability to succeed in medical study with the same accuracy as subject GPAs. Advantage of overall GPA is that it describes the breadth of student's knowledge and that it should be less affected by the teacher effect. Profile subjects nevertheless are of some importance: condition A2 was found to be important.

Optimal model contained AT, GPA, conditions A1 and A2 and accounted for 30% of variability of predicted value. In this sense we have also proved suitability of current admission process. Instead of weighted sum of optimal predictors, the school offers places to given number of applicants with best GPA and satisfying conditions A1, A2 and A3,

while the rest of the students selected solely on the basis of AT. As an advantage of this two-step process, 1st LF UK saves money and also encourages talented students to study at 1st LF UK, particularly those who would not have even attempted to pass the difficult and memorizing admission test. As a disadvantage, the two-step process is less accurate, since either only GPA or only AT is used for selection of students but never both together. As a consequence, possibilities of future analyses are limited due to data incompleteness. Incorporating GPA as weighted score for the group of students who also took AT should bring improvement of the selection process and more complete data.

In most of the studies concerning admission process, the study success comes from practical reasons defined only by average GPA in the first years of medical study. Höschl and Kožený (1997) looked for optimal predictor of academic performance in the first three years of medical study and found associations in higher class of medical study to be weaker. They explained this feature by the fact that in the first two years, the contact with patients is still limited and the processing of theoretical knowledge constitutes the main criterion for evaluating students' performance. They concluded that admission committees are interested in predicting probability of success or failure rather than estimating grade point averages, which are primary of theoretical and illustrative significance only. In our study, we fulfil recommendations given in mentioned work by analysing the overall success in medical study. A great disadvantage of this approach is the fact that we had to wait more than seven years; meanwhile the population of admitted students might have changed as well as curriculum or admission criteria.

Optimal model for the prediction of study success in Höschl and Kožený (1997) contained high-school GPA in physics and the admission subtest in physics. Kožený and Tišanská (2001) found mathematics to be the best high-school subject to predict success in medical study. Both physics and mathematics involve logical thinking more than memorizing. In our study, the difference of these two subjects from other subjects is proved by cluster analysis. The admission test in physics was moreover the only one whose removal from the admission test would cause significant loss of information. We assume, that subtest in physics would have differed even more from other two subtests if the most of the questions (or their analogous versions) were not published in advance. Published

items cause the current admission test to 1st LF UK to be the test of willingness and ability to memorize large amount of information.

Mentioned qualities are in view of further medical study necessary due to huge volume of curriculum, and our analyses prove that they are of use at 1st LF UK. Nevertheless, evolution of society and progress of information technologies in the last twenty years raise doubts, whether the paradigm is still true. Strategy of large amount of published items resulted from requirements on transparency of the admission process and equal chances for all applicants. By oversatisfying these requirements we lose the possibility to test logical thinking and ability to creatively apply the gained knowledge, which should be crucial in further clinical praxis of medical graduates.

CONCLUSION

In spite of differences in the level, degree and standards secondary schools of particular type may have, preadmission grades (GPA) were shown to predict overall success in medical study almost with the same accuracy (and even a bit higher) as admission tests (AT). While GPA accounted for about 15% of variability of predicted value, single AT accounted for only 13%. Overall GPA was shown to have similar predictive power as optimal combination of subject GPAs. Further it shows up that equal weighting of the three admission subtests is optimal; nevertheless, chemistry and biology subtests may be removed without significant loss of information.

AT and GPA were shown to describe different aspects of students' abilities to successfully graduate from medical school.

Incremental validity of admission tests over preadmission grades was proved; optimal combination of AT and GPA explained 22% of variability of success. Adding information on year of graduation and type of high school (which together with GPA strongly influences fulfilment of conditions on profile classes) raised percentage of explained variability to 31%.

Current admission criteria to 1st LF UK are in agreement with findings of this study. Offering places without AT to students with best pre-admission GPA is reasonable as well as additional criteria on profile subjects and year of graduation. Better results might be reached by incorporating GPA into admission process of all students.

References:

Carstensen, B. et. al. (2009). Epi: A package for statistical analysis in epidemiology. R package version 1.0.12. URL <http://www.pubhealth.ku.dk/~bxc/Epi/>

Emery J.L., Bell J.F. (2009). The predictive validity of the BioMedical Admissions Test for pre-clinical examination performance. *Medical Education* 43(6), pp. 557-564.

Harrell, F.E. Jr., et al (2012). rms: Regression Modeling Strategies. R package version 3.5-0. URL <http://CRAN.R-project.org/package=rms>

Höschl, C., Kožený J. (1997). Predicting academic performance of medical students: the first three years. *American Journal of Psychiatry*, 154(6 Suppl.), pp. 87-92.

Kožený J., Tišanská L., Höschl C. (2001). Academic success at the high school: predictor for graduation at medical faculty [Akademická úspěšnost na střední škole: prediktor absolvování studia medicíny, in Czech], *Československá psychologie*, 45, 1, 2001, pp. 1-6.

McMannus I.C. et al. (2011). Predictive validity of the Biomedical Admissions Test: an evaluation and case study. *Medical Teacher* 33(1), pp. 53-57.

Parry, J., Mathers J., et al. (2006). Admissions processes for five year medical courses at English schools: review. *British Medical Journal* 332, pp. 1005-1009.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>

Salvatori, P. (2001). Reliability and validity of admissions tools used to select students for the health professions. *Advances in Health Science*, 6, pp. 159–175.

Línová, M. (2012). Přehled o průběhu přijímacího řízení pro akademický rok 2002/2003 - 2010/2012, in Czech, URL <http://www.lfl.cuni.cz/prehled-o-prubehu-prijimaciho-rizeni>

Štuka, Č., Šimeček, P. (2006). Studium souvislosti mezi úspěšností studia medicíny, známkami studentů na střední škole a výsledky přijímacích zkoušek, *Sborník Medsoft 2006*.

Štuka, Č., Smutek, D. (2003). Correlation of the Entrance Examination Result and the Study Success at the Medical Faculty, Charles University Prague. International Symposium on

Information and Communication Technologies ISICT03, Dublin 2003, *Proceedings of ISICT 2003*, p. 52 - 57, ISBN 0-9544145-2-7, URL <http://dl.acm.org/citation.cfm?id=963600.963610>

Venables, W. N., Ripley, B. D. (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Predictive Validity of Scientific Knowledge Test and Intellectual Aptitude Test Used in Admission for Medical School Performance during the First Year

Cestmir Stuka¹, Patricia Martinkova², Marek Novak¹, Karel Zvara², Stanislav Stipek¹

¹ First Faculty of Medicine of Charles University in Prague, Czech Republic

² Centre of Biomedical Informatics, Department of Medical Informatics, Institute of Computer Sciences of the Academy of Sciences of the Czech Republic, Prague

Abstract

Introduction: General aptitude test (GAT) was temporarily added to admission process at the First Faculty of Medicine of Charles University in Prague (FFM).

Aims: To prove predictive validity of GAT and to compare it with validity of current scientific knowledge test (ST). The added value of GAT was also examined.

Methods: Prediction validity was estimated by Pearson correlation within cohorts (years 2002 – 2009) and by the correlation after each variable has been group-mean centred for all cohorts together. Incremental validity was assessed by testing submodels in linear and logistic regression models.

Results: The first year of studies is shown to be crucial for overall academic performance at FFM. Although the predictive validity of GAT is significantly nonzero; it is markedly lower than predictive validity of ST. Also, ST contributes significantly to the proper selection of students, whereas the contribution of GAT can be inferred only in part.

Conclusion: We found similar relations concerning ST and GAT regardless if they were observed in the Czech Republic or the United Kingdom.

Introduction

The selection of students for medical studies is an important task and it receives a constant attention. Information on academic performance at secondary schools, entrance exams, interviews and other methods are used in admission process.

The purpose of the admission procedure is to select students who will be able to successfully complete their entire studies. Faulty selection leads to the ineffective use of the teaching capacity of a school on the one hand, and to the loss of time of students spent by studying in an unsuitable field on the other hand. In terms of schools, the losses resulting from an incorrect choice are directly related to an economic loss. From the perspective of students, a properly set admission procedure could help applicants determine whether they have the necessary aptitude and ability for the demanding studies (Kuncel, Ones et al. 2001).

Admission procedures at Czech faculties of medicine

All seven Czech faculties of medicine use some kind of scientific knowledge tests (ST) containing items in physics (P), chemistry (C) and biology (B) in the admission procedure. Some schools supplement the tests with an interview, interpretation of a text or a general aptitude test (GAT), which is a test of intellectual aptitude similar to an abbreviated IQ test.

In 2002, First Faculty of Medicine of Charles University in Prague (FFM), and in subsequent years several other Czech medical schools, implemented the possibility of admissions without having to take entrance exams but only based on outstanding undergraduate grade point averages (uGPAs). Information about admission procedures for the year 2010 at faculties of medicine in the Czech Republic is summarized in Table 1.

Table 1: Admission procedures at faculties of medicine in the Czech Republic in 2010

School	Test in Biology, Chemistry, Physics	Sample items known in advance	General Aptitude Test (GAT)	Interview	Interpretation of a text	Admission based on uGPA	Number of admitted*	Number of applicants*
First Faculty of Medicine of Charles University in Prague	1	1	1	0	0	1	450	3,374
Second Faculty of Medicine of Charles University in Prague	1	1	1	1	0	0	150	1,813
Third Faculty of Medicine of Charles University in Prague	1	0	0	1	1	1	155	1,413
Faculty of Medicine of Charles University in Hradec Králové	1	1	0	0	0	1	165	1,415
Faculty of Medicine of Charles University in Pilsen	1	1	0	0	0	1	230	1,850
Faculty of Medicine of Palacký University in Olomouc	1	1	0	0	0	1	220	1,838
Faculty of medicine of Masaryk University in Brno	1	1	0	0	0	0	400	2,894

*The number of admitted students and applicants applies to study of general medicine in the national language.

Admission procedure at the First Faculty of Medicine in Prague

FFM is the largest of seven Czech faculties of medicine. Each year approximately 450 students, out of usually 3,500 applicants, are admitted to study general medicine in the national language at this faculty.

Multiple choice tests from three scientific subjects have been used in admission process at FFM since 1992. Before this, the admission procedure included interviews, which since 1992 have been omitted from exams in order to eliminate all personal influences. Items used in knowledge tests are created in scientific institutes of FFM. Sample items are made public in advance in order to ensure equal access to study materials.

In 2008, General Aptitude Test (GAT), with questions similar to those used in IQ tests, was experimentally added to ST. The GAT was developed at the Faculty of Arts of Charles University in Prague. Unlike the ST measuring the knowledge of chemistry, biology and physics, the new test focused on intellectual aptitude, particularly on logical and analytical skills and spatial reasoning.

The appropriateness of intellectual aptitude test was repeatedly questioned:

Both McManus and Emery (Emery and Bell 2009) (McManus, Ferguson et al. 2011) point out the different contributions of knowledge and aptitude tests in Biomedical Admissions Tests (BMATs) as predictors of academic performance. In their works authors show that in BMAT, used in the UK, its Section 2, focusing on Scientific Knowledge and Applications, correlates clearly better to grade point averages in the first and second year of studies than Section 1 focusing on Aptitude and Skills. In their other work (Emery, Bell et al. 2011) authors further show that in order to predict the grade point average in the first year of studies at a university, Section 2 significantly contributes to the prediction while Section 1 does not.

The appropriateness of the use of intelligence tests for the selection of candidates for studying medicine in UK was discussed by McManus et al. (McManus, Powis et al. 2005). They summarise previous findings (McManus, Smithers et al. 2003) and the conclusions of other studies stating that “Since the 1970s, university achievement has been shown to be predicted by A levels but not by intelligence tests” and that “Schools are introducing tests of intellectual aptitude without evidence of appropriateness, accuracy, or added value, making them open to legal challenge”.

Aims

The primary objective of this work is to compare the predictive validity of scientific knowledge test (ST) and aptitude test (GAT) as parts of admission procedure for medical study at FFM, and to decide, whether the two parts significantly contribute to the prediction of the success of students' study.

With reference to relevant works in UK we try to generalize our findings and to discuss the appropriateness of both types of admission tests, which measure the different skills and abilities of applicants.

Methods

Admission tests

Since 1992 the admission procedure at FFM consists of written scientific knowledge tests (ST). In order to ensure the equal chances of applicants, it was decided to make public approximately 1,300 sample items for each scientific subtest (P, C, B) in advance. The admission test then comprises these questions themselves or with a slight variation (e.g. changes in the numerical values of examples or a change in the order of answers).

Admission test (AT) always has a total of 300 items. Until 2007 it was 100 randomly selected multiple-choice items for each subtest of ST (Physics, Biology and Chemistry). Since 2008, after the addition of GAT, the number of questions on knowledge tests was reduced to 240 (80 for each subtest) and the GAT had 60 questions.

The test items allow to choose one or more correct answers from four possible answers. A correct item is considered to be the one in which all of the correct answers are marked.

The points acquired from individual subtests are added together and based on this the final ranking of candidates is determined. Then, according to this ranking, the determined number of the best applicants is admitted so as to fully utilize the capacity of a school, i.e. in our case around 450 students. This number and admission test method applies to students studying in the national language. Students studying concurrently in English are counted separately and their admission procedure is different.

In 2002, FFM implemented the possibility of admitting some students based on their undergraduate grade point averages (uGPAs). The percentage of students admitted in this manner gradually increased from 25% to the current approximately 33% of all admitted students. We discuss this in detail with the comparison of the academic performance of students admitted in both ways in a separate article (Martinkova, Stuka et al.). The predictive abilities of admission tests (AT) and uGPAs have shown to be essentially comparable. However, it seems that each method captures a different part of an applicant's abilities. A combination of both (AT and uGPA) would show significantly greater predictive validity than either method alone.

However, the comparison of grades at secondary schools is only possible at comparable schools. Although most students come from Gymnasiums, whose difficulty and curriculums are comparable, some applicants come from technical secondary schools, which have different curriculums and levels of difficulty. Although there is an option of implementing correction factors for the correction of uGPAs from technical secondary schools, the values of correction factors could be the subject of disputes that are difficult to resolve. Instead, supplementary criteria on students who want to be admitted without AT based on their uGPAs were implemented (the completion of specific subjects to a given extent, etc.) that de facto define the type of school. Admission tests serve to assess students' abilities regardless of the type of secondary school they attended.

In 2008, General aptitude test (GAT) was included in the set of admission tests at FFM - an intellectual aptitude test, whose questions are similar to those used in IQ tests. It comprises of 60 questions focusing on conceptual, spatial and numerical reasoning and comes from the Department of Psychology, the Faculty of Arts of Charles University in Prague. In the first year (2008), the GAT was added as a trial - it was not used for the ranking of applicants. In 2009, the GAT was used for the selection of candidates for the first time.

Academic performance

Studying medicine at faculties of medicine in the Czech Republic lasts six years. However, during their studies many students use the opportunity to study abroad for one or two semesters. Because in the assessment of tests it would be impractical to wait for the completion of studies, an alternate criterion is used against which is examined the predictive ability of admission tests - mostly the grade point average in the first year of studies. In our case this premise is justified because the first years of studies are very difficult: Figure 1 depicts the percentage of students who dropped out of the FFM in each year of the study for students admitted in 2002 and 2003. Graduates are depicted in the upper part and unsuccessful students in the lower part of the graph. In both years the percentage of unsuccessful students (shown by the horizontal dividing line) is around 40%.

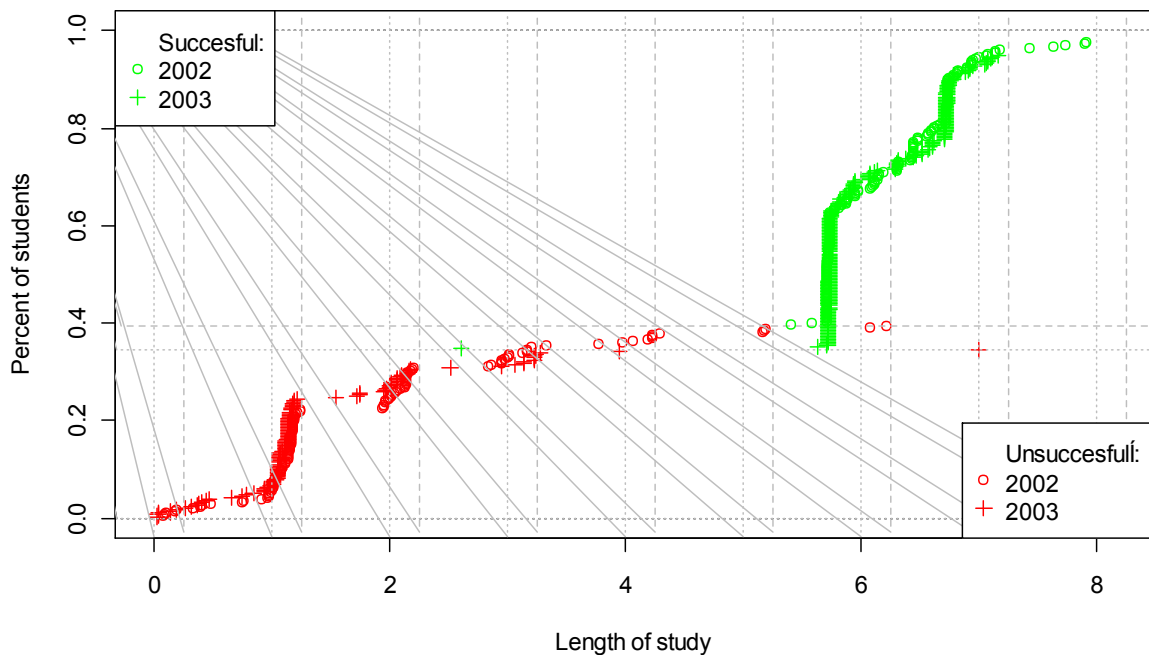


Figure 1: Length of studies of successful and unsuccessful students admitted in 2002 and 2003

The course of the leaving study shows the difficulty of the first year. Most students who drop out after failing do so within three months after the first year. Due to its high difficulty level, the first year at this school can be perceived as part of the admission procedure (Sulek and Hilser 2008). If the curriculum would change in order to avoid the said phenomenon, it would be necessary to increase the capacity of the faculty in higher (clinical) years so that it would be able to handle the non-reduced number of students.

Success in the first year is defined using information as to whether a student did not drop out within 3 months of the official end of the first academic year, i.e. by December, 31st. Formally, the decisive date for measuring success should be October, 30th, because the official end of the exam period ends on September, 30th, and a month is the statutory period for handling an appeal, but in practice many appeals take place and if we wait until the end of a relevant calendar year, we acquire much more accurate information as to whether a student passed the year or not (see Figure 1).

For a finer assessment of the success, we also use the **overall grade point average in the first year** (GPA1). The required subjects are the same for all students during the first year of study and over the years they changed very little. Required subjects included Biophysics, Latin, Medical Informatics, Histology and Embryology, Anatomy, and a Foreign Language. A student on an exam is graded with the grades of 1 (excellent) to 4 (failed). A student has the option of taking one regular examination and two re-sittings of an examination of a given subject. The average was calculated from all attempts, and if a student did not come to an examination, he received a 3 times failed grade for the purposes of this analysis.

Data set

In this paper, we focus on applicants admitted to FFM in the years 2002 to 2009. We included only applicants applying to the field of general medicine in Czech language, who applied for admission to this faculty for the first time. In Table 2, we can see the number of students who started the FFM in years 2002 – 2009 and were admitted based on admission tests and based on uGPAs. For analysis in this work, only those who passed admission tests were used.

Table 2: Number of students who were admitted to the field of general medicine taught in Czech language at FFM and started the study

Year	2002	2003	2004	2005	2006	2007	2008	2009
Admitted via a test	276	277	280	286	323	214	261	311
Admitted via uGPAs	94	96	142	100	99	99	115	155

The differences in the number of students in cohorts are due to changes in the admission policies of FFM, for example, the varying number of students admitted solely on the basis of their uGPA, as well as changes in other faculties (the expanding of admission based on uGPA led to the multiple applications of the best students and an increase in the number of students that finally did not start studies at a given faculty)

The data set contains, among other, the following data about students concerning their success in studies:

Current study status; date when study status came into being,

Academic year of admission

The grade point average of a given year (all general medicine students in a given cohort have the same prescribed tests in their first years of studies),

Number of re-sitting examinations of subjects from which the grade point average is calculated (each examination has one regular examination and two re-sittings of an examination, therefore a total of 3 attempts)

Data on *the number of points* earned in each of four tests of the admission procedure (*ST* and its subtests in *Biology, Chemistry and Physics*, and for cohorts 2008 and 2009 also *GAT*).

Statistical analysis of data

Predictive validity of admission tests generally shows to what extent a given variable is able to predict future success of a student in the university study. We assess the validity of admission tests as the degree of dependence between individual admission subtests and the grade point average during the first year of study (GPA1). We use *Pearson's correlation coefficient* for measuring dependence in individual cohorts. For the overall assessment of multiple cohorts, we calculate the correlation between admission tests and the grade point average after each variable has been group-mean centred (Robinson 1950).

We assess incremental validity using testing submodels in regression models. We use a linear regression model for modelling the dependence of GPA1, and we use logistic regression for modelling the dependence of success in the first year (dichotomous variable dropped out / continued on).

Statistical analysis was performed and figures were plotted in free statistical environment *R* (R_Development_Core_Team 2009), including its libraries *multilevel* (Bliese 2008) and *meta* (Schwarzer 2010).

Results

The values of the correlation coefficients between GPA1 and admission tests (GAT, ST and its subtests C, P and B) are shown in Figure 2. We can conclude that the validity of the tests varies within individual years. All correlations are significantly nonzero. Also, there is a significant correlation for the GAT, but it is clearly the smallest.

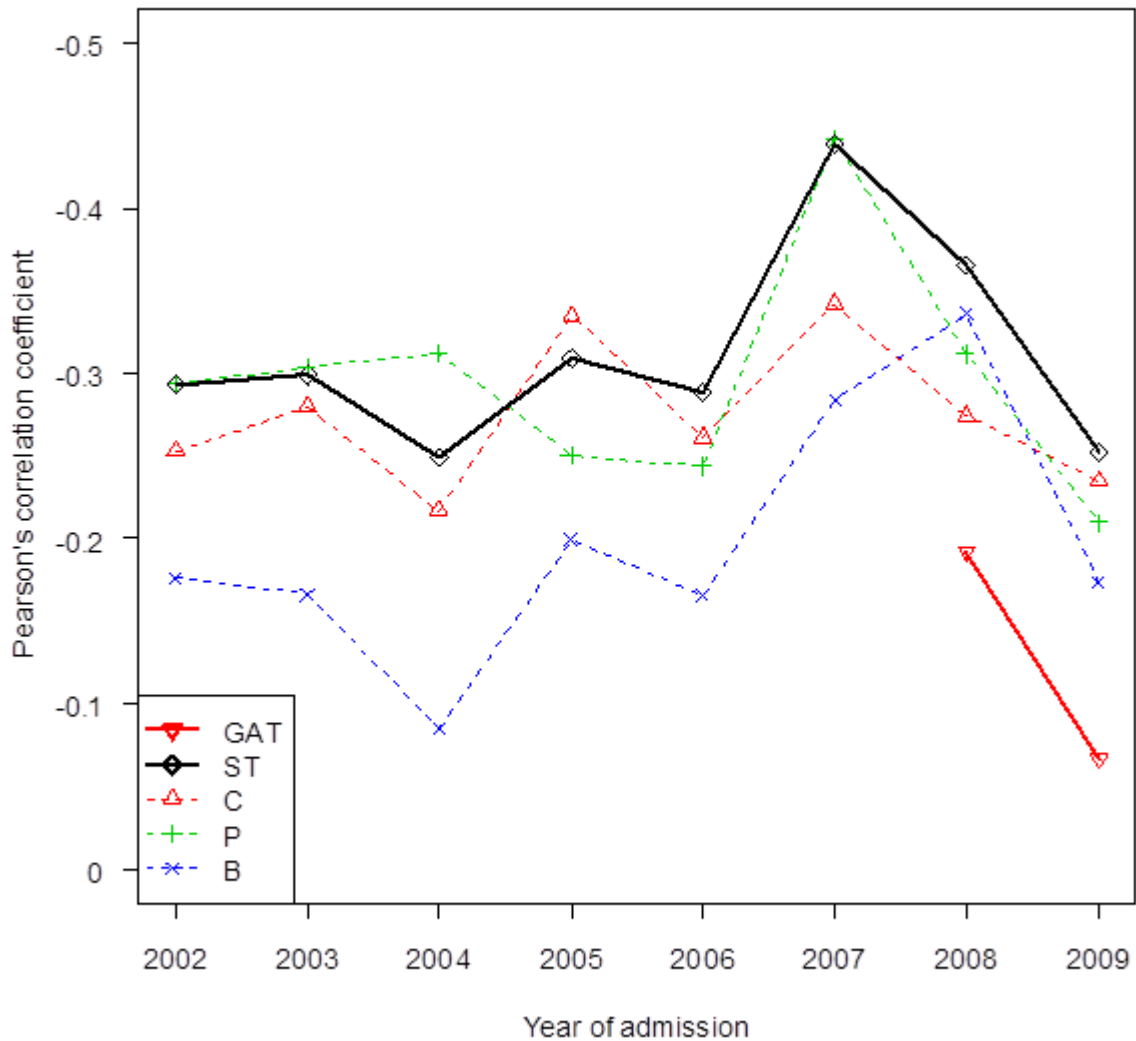


Figure 2: Validity of individual parts of the admission test for individual cohorts (years of admission) ST - Scientific knowledge test, C – Chemistry test, P – Physics test, B – Biology test, GAT - General aptitude test (IQ)

In Table 3, the information from Figure 2 is summarized by computing the correlation of AT and GPA1 after each variable has been group-mean centred. The first column shows correlations calculated from data of those admitted in the years 2002-2009). The second column show correlation calculated solely from the years 2008-2009, for which we also have available data from GAT. We got almost identical values using a meta-analysis (Schwarzer 2010).

Table 3: Correlations of admission tests and GPA1 after each variable has been group-mean centred.

Admission test	Years 2002-2009	Years 2008-2009
ST	-0.30	-0.30
Subsets:		
C	-0.27	-0.25
P	-0.29	-0.26
B	-0.19*	-0.25
GAT		-0.13

ST - Scientific knowledge test, C – Chemistry test, P – Physics test, B – Biology test, GAT - General aptitude test (IQ)

**test from Biology, as is evident from Figure 2, shows poor results especially in earlier years*

Besides predictive validity, also incremental validity of tests was studied. For assessing the incremental validity of the newly added GAT, we used only the cohorts admitted in the years 2008 and 2009, for which we have GAT data available. When predicting **GPA1** by ST and by the newly implemented GAT, both of them significantly ($p < 0.001$) contribute to the prediction. When predicting **success in the first year of study** (dichotomous variable dropped out / continued on) only ST contribute significantly ($p < 0.001$) and effect of GAT was no longer significant ($p = 0.286$).

Discussion

We noted that the substitution of the criteria of success of the entire length of studies with the criteria of GPA1 of studies is acceptable especially because we know (Figure 1) that success in the first year, in most cases, is decisive for the entire length of studies.

We proved that all parts of scientific knowledge test (Biology, Chemistry and Physics) are beneficial for the proper selection of students. In comparison, the predictive validity of the new GAT is lower and its contribution can be inferred only partially (it significantly adds new information if we predict GPA1, however, if we predict actual success in the first year of studies, its contribution is not significant).

It turns out that the predictive validity of individual tests varies between years. This is due to the established methodology of the (random) selection of test items from the battery of items, without taking into account their quality, ascertained via an item analysis. In 2007, we conducted an item analysis of the first round of admission tests that took place in 1999. The time interval of eight years, during which the admitted students had enough time to complete their studies, enabled us to assess the actual validity of items (not only the validity relative to the first year of studies). This item analysis showed that the quality of questions varies widely from explicitly useful questions to questions that correlated negatively with the entire results of studies, i.e. students who were unsuccessful in overall studies were more likely to choose the right answer to the question (Stuka, Martínková et al.) (to be published).

If we compare the predictive validity of individual subtests, we see that the quality of knowledge tests is comparable, although it greatly varies from year to year. The contribution of the general aptitude test for the years 2008-2009 is considerably smaller.

Similar relations between tests were also found in schools in the UK (McManus, Powis et al. 2005). Therefore, it seems that when assessing the appropriateness of tests, we find similar relations regardless if they were observed in the Czech Republic or the United Kingdom.

References

- Bliese, P. (2008). Multilevel Functions, R package version 2.3.
- Emery, J. L. and Bell J. F. (2009). "The predictive validity of the BioMedical Admissions Test for pre-clinical examination performance." Med Educ **43**(6): 557-564.
- Emery, J. L., Bell J. F., et al. (2011). "The BioMedical Admissions Test for medical student selection: issues of fairness and bias." Medical Teacher **33**(1): 62-71.
- Kuncel, N. R., Ones D. S., et al. (2001). "A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance." Psychological Bulletin **127**(1): 162-181.
- Martinkova, P., Stuka C., et al. Analysis of current admission criteria for First Faculty of Medicine at Charles University in Prague. (to be published)
- McManus, I. C., Ferguson E., et al. (2011). "Predictive validity of the Biomedical Admissions Test: an evaluation and case study." Med Teach **33**(1): 53-57.
- McManus, I. C., Powis D. A., et al. (2005). "Intellectual aptitude tests and A levels for selecting UK school leaver entrants for medical school." British Medical Journal **331**(7516): 555-559.
- McManus, I. C., Smithers E., et al. (2003). "A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study." British Medical Journal **327**(7407): 139-142.
- R Development Core Team (2009). R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.
- Robinson, W. S. (1950). "Ecological correlations and the behaviour of individuals." American Sociological Review **15**(3): 351-357.
- Schwarzer, G. (2010). Meta-Analysis with R. R package version 1.6-1
- Stuka, C., Martinkova P., et al. Item Analysis of Medical School Admissions Test, First Faculty of Medicine, Charles University in Prague (to be published).
- Sulek, S. and Hilser M. (2008). Strong and weak points in the general medicine course education at the First Faculty of Medicine in Prague: students' view. AMEE 2008. Prague.

Závěr

V literární rešerši jsme se zabývali metodami výběru uchazečů o vysokoškolské studium se zvláštním přihlédnutím k výběru studentů medicíny. Kromě závěrů, které jsme již zmínili v příslušné kapitole stojí za vyzdvižení rozdíl, mezi znalostními testy jak jsou používány na 1. LF UK v Praze a znalostními testy známými např. z Velké Británie (scientific knowledge test, součást BMAT). V prvním případě jsou totiž testové otázky předem známé a často i formulované tak, že se testovaný jedinec odpovědi raději naučí. V zahraničních testech typu „scientific knowledge test“ se netestuje samotná znalost, ale spíše schopnost znalosti uplatnit. To je podstatný rozdíl, který omezuje možnost aplikovat zákonitosti zjištěné v zahraničí na naše poměry.

Otevřeným problémem testování uchazečů o studium na 1. LF UK je rovněž absence měření kvality testových otázek. To je ve srovnání s metodikami používanými v zahraničí poměrně podstatná anomálie, kterou bude třeba řešit.

V předběžné studii jsme probrali používané metodické postupy a ukázali jsme, že při hodnocení výkonu na střední škole, lze využít kumulativní informaci o středoškolských průměrech, místo detailní znalosti známek z jednotlivých předmětů, což nám otevírá cestu pro hodnocení ročníků, kde detailní informace již chybí.

V další části práce jsme na datech studentů přijatých ke studiu v roce 1999 studovali jednotlivé prediktory celkového studijního úspěchu a prokázali jsme, že studijní průměry ze střední školy jsou stejně dobrým prediktorem jako výsledky znalostních přijímacích testů a přitom každá z těchto veličin popisuje jinou dimenzi studentových schopností. Potenciální výhoda plynoucí z použití obou metrik současně je evidentní.

V poslední části analyzujeme přínos nově zavedeného testu Všeobecné studijní připravenosti ve vztahu k dříve zavedeným znalostním testům. Prokázali jsme, všechny části znalostního testu (B, CH, F) jsou přínosem pro správný výběr studentů. Přínos testu VSP je naproti tomu sporný. Je přínosem pokud predikujeme v prvním ročníku studijní průměr, ale není signifikantní, pokud predikujeme ve stejném ročníku studijní úspěch (prošel / neprošel).

Do budoucna bychom se chtěli zabývat i položkovou analýzu testových otázek a možností využití nově zaváděných státních maturit jako jednoho z prediktorů.

Literatura:

- ACER. (2011, 2011). "Undergraduate Medicine and Health Sciences Admission Test (UMAT)." 2011, from http://umatweb.acer.edu.au/images/infobook/UMAT_InfoBook.pdf.
- Albanese, M. A., T. H. Kent, et al. (1977). "A comparison of the difficulty, reliability and validity of complex multiple choice, multiple response and multiple true-false items." *Annu Conf Res Med Educ* **16**: 105-110.
- Albanese, M. A., T. H. Kent, et al. (1979). "Cluing in multiple-choice test items with combinations of correct responses." *J Med Educ* **54**(12): 948-950.
- Anděl, M. (2010). Informace o studiu a přijímacím řízení pro akademický rok 2011/2012. Praha, Univerzita Karlova v Praze, 3. lékařská fakulta.
- Anděl, M. (2010). Výroční zpráva 3. lékařské fakulty za rok 2010. Praha, Univerzita Karlova v Praze, 3. lékařská fakulta.
- Anděl, M., C. Höschl, et al. (1993). "Jak se stane ze studenta medicíny dobrý lékař." *Vesmír* **72**(5).
- Beaujean, A. A., M. W. Firmin, et al. (2006). "Validation of the Frey and Detterman (2004) IQ prediction equations using the Reynolds Intellectual Assessment Scales." *Personality and individual differences* **41**(2): 353-357.
- Beller, M. (1992). "Psychometric, Social and Political Issues in Admissions to Higher-Education in Israel." *International Journal of Psychology* **27**(3-4): 569-569.
- BenShakhar, G., I. Kiderman, et al. (1996). "Comparing the utility of two procedures for admitting students to liberal arts: An application of decision-theoretic models." *Educational and Psychological Measurement* **56**(1): 90-107.
- Bridgeman, B., L. McCamley-Jenkins, et al. (2000). Predictions of Freshman Grade-Point Average From the Revised and Recentered SAT I: Reasoning Test. *College Board Research Report No. 2000-1*. New York, College Entrance Examination Board.
- Callahan, C. A., M. Hojat, et al. (2010). "The predictive validity of three versions of the MCAT in relation to performance in medical school, residency, and licensing examinations: a longitudinal study of 36 classes of Jefferson Medical College." *Acad Med* **85**(6): 980-987.
- Clare, J. (1999). SAT – a simple way to grade students. *The Daily Telegraph*.
- Coates, H. (2008). "Establishing the criterion validity of the Graduate Medical School Admissions Test (GAMSAT)." *Med Educ* **42**(10): 999-1006.
- collegeboard.org. (2011). "What is the SAT?" Retrieved 6.7.2011, 2011, from <http://professionals.collegeboard.com/testing/sat>.
- Coumarbatch, J., L. Robinson, et al. (2010). "Strategies for Identifying Students at Risk for USMLE Step 1 Failure." *Family Medicine* **42**(2): 105-110.
- Didier, T., C. D. Kreiter, et al. (2006). "Investigating the utility of a GPA institutional adjustment index." *Advances in Health Sciences Education* **11**(2): 145-153.
- DomeSAT. (2009). "SAT Test Demographics by Income and Ethnicity." from <http://www.domesatreview.com/content/sat-test-demographics-income-and-ethnicity>.
- Donnon, T., E. O. Paolucci, et al. (2007). "The predictive validity of the MCAT for medical school performance and medical board licensing examinations: a meta-analysis of the published research." *Acad Med* **82**(1): 100-106.
- Downing, S. A. (2002). "Threats to the validity of locally developed multiple-choice tests in medical education: Construct-irrelevant variance and construct underrepresentation." *Advances in Health Sciences Education* **7**(3): 235-241.

- Dušek, L. (2004). Lékařská fakulta MU: Analýza výsledků přijímacích zkoušek 2003-4. Brno, Institut biostatistiky a analýz, Masarykova univerzita: 39.
- Emery, J. L. and J. F. Bell (2009). "The predictive validity of the BioMedical Admissions Test for pre-clinical examination performance." Med Educ **43**(6): 557-564.
- Emery, J. L., J. F. Bell, et al. (2011). "The BioMedical Admissions Test for medical student selection: issues of fairness and bias." Medical Teacher **33**(1): 62-71.
- Fernando, N., G. Prescott, et al. (2009). A comparison of the United Kingdom Clinical Aptitude Test (UK-CAT) with a traditional admission selection process. Medical Teacher, Taylor & Francis Ltd. **31**: 1018-1023.
- Frey, M. C. and D. K. Detterman (2004). "Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability (vol 15, pg 373, 2004)." Psychological science **15**(9): 641-641.
- Gilbert, G. E., W. T. Basco, Jr., et al. (2002). "Predictive validity of the Medical College Admissions Test Writing Sample for the United States medical licensing examination steps 1 and 2." Adv Health Sci Educ Theory Pract **7**(3): 191-200.
- Hach, P. (2010). Od kdy máme přijímací testy. Č. Štuka. Praha, 1. LF UK Praha 1.
- Hach, P., J. Kraml, et al. (1994). "Vědomosti, inteligence, osobnost a ars medici." Vesmír **73**(33): 2.
- Hall, F. R. and B. A. Bailey (1992). "Correlating students' undergraduate science GPAs, their MCAT scores, and the academic caliber of their undergraduate colleges with their first-year academic performances across five classes at Dartmouth Medical School." Acad Med **67**(2): 121-123.
- Havlová, M. (2010). "Informace o studiu a přijímacím řízení." from <http://www.lf2.cuni.cz/Studium/pr/informace.htm>.
- Hojat, M., B. D. Borenstein, et al. (1988). "Cognitive and noncognitive factors in predicting the clinical performance of medical school graduates." J Med Educ **63**(4): 323-325.
- Hojat, M., M. Robeson, et al. (1993). "Students Psychosocial Characteristics as Predictors of Academic-Performance in Medical-School." Academic Medicine **68**(8): 635-637.
- Hojat, M., W. H. Vogel, et al. (1988). "Effects of Academic and Psychosocial Predictors of Performance in Medical-School on Coefficients of Determination." Psychological Reports **63**(2): 383-394.
- Jones, L. V. (1994). "Perspectives on Educational Testing: Discussion." Educational Measurement: Issues and Practice **13**(2): 28-30.
- Julian, E. R. (2005). "Validity of the Medical College Admission Test for predicting medical school performance." Acad Med **80**(10): 910-917.
- Kasal, P. (1997). "Problémy spojené se "zaškrťávacími testy" ve výuce." Pelikán - Akademický bulletin 2. lékařské fakulty.
- Kennet-Cohen, T., S. Bronner, et al. (1999). The predictive validity of the components of the process of selection of candidates for higher education in Israel, National Institute of Testing and Evaluation (NITE): 25.
- Kobrin, J. L., B. F. Patterson, et al. (2008). Validity of the SAT for Predicting First -Year College Grade Point Average. Research Report No. 2008-5. New York, College Board. **5**: 10.
- Kreiter, C. D. and Y. Kreiter (2007). "A validity generalization perspective on the ability of undergraduate GPA and the medical college admission test to predict important outcomes." Teach Learn Med **19**(2): 95-100.
- Lang, D. M. (2007). "Class Rank, GPA, and Valedictorians: How High Schools Rank Students." American Secondary Education **35**(2): 12.

- Lech, J. C. (2007). The math-verbal achievement gap: Are we getting better at math, or are the SAT questions getting easier?, Peabody College for Teachers of Vanderbilt University.
- Línová, M. (2011). "Přehled o průběhu přijímacího řízení pro akademický rok 2002/2003 - 2010/2011." from <http://www.lf1.cuni.cz/prehled-o-prubehu-prijimaciho-rizeni>.
- Lynch, B., R. MacKenzie, et al. (2009). "Does the UKCAT predict Year 1 performance in medical school?" Medical Education **43**(12): 1203-1209.
- Margolis, M. J., B. E. Clauser, et al. (2010). "Validity evidence for USMLE examination cut scores: results of a large-scale survey." Acad Med **85**(10 Suppl): S93-97.
- Martínková, P., K. Zvára, et al. (2009). Analýza zápisů studentů přijatých v roce 2007 na 1. LF. Praha, 1. LF UK Praha: 10.
- McCoubrie, P. and L. McKnight (2008). "Single best answer MCQs: a new format for the FRCR part 2a exam." Clinical Radiology **63**(5): 506-510.
- McDonald, A. S., P. E. Newton, et al. (2001) "Aptitude Testing for University Entrance." A Literature Review.
- McManus, I. C., E. Ferguson, et al. (2011). "Predictive validity of the Biomedical Admissions Test: an evaluation and case study." Med Teach **33**(1): 53-57.
- McManus, I. C., D. A. Powis, et al. (2005). "Intellectual aptitude tests and A levels for selecting UK school leaver entrants for medical school." British Medical Journal **331**(7516): 555-559.
- McManus, I. C., E. Smithers, et al. (2003). "A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study." British Medical Journal **327**(7407): 139-142.
- MEDICS, G. (2005). "Launching medical careers." 2011, from <http://www.gapmedics.com>.
- Přenosi, J. (2004, 2004). "Přijímání studentů k řádnému studiu na Univerzitě Karlově." iForum, Z historie UK, 2004, from <http://iforum.cuni.cz/IFORUM-471.html>.
- Refor, C. f. E. (1996). SAT Increase -- The Real Story, Part II Washington, DC, The Center for Education Refor.
- Rothstein, R. (2002). Sums vs. Summarizing: SAT's Math-Verbal Gap. The New York Times. NY.
- Schneider, D. and N. Dorans (1999). Concordance Between SAT I and ACT Scores for Individual Students. New York, College Boar.
- Slezák, R. (2009). Výroční zpráva 2009, LF KH, Univerzita Karlova v Praze, Lékařská fakulta v Hradci Králové.
- Štuka, Č. and J. Klaschka (2003). Statistické vyhodnocení přijímacího řízení a výsledků ze střední školy ve vztahu k úspěšnosti studia na vysoké škole. XXVI. dny lékařské biofyziky. Senohraby, Univerzita Karlova Praha, 1. lékařská fakulta. **XXVI.**
- Topičová, V. (2011). Tabulka k přijímacímu řízení na lékařských fakultách v ČR v roce 2010. Č. Štuka. Praha, Studijní odd. 1.LF UK v Praze.
- UKMS. (2011). "UK Medical Schools." 2011, from <http://www.ukmedicalschoools.com>.
- Vejražka, M. (2011). Tvorba otázek přijímacích testů. Č. Štuka. Praha.
- Wainer, H. (1999). "Comparing the incomparable: an essay on the importance of big assumptions and scant evidence." Educational Measurement: Issues and Practice **18**(4): 6.
- Yates, J. and D. James (2010). "The value of the UK Clinical Aptitude Test in predicting pre-clinical performance: a prospective cohort study at Nottingham Medical School." BMC Med Educ **10**: 55.
- Zeidner, M. (1986). "Are Scholastic Aptitude-Tests in Israel Biased Towards Arab College-Student Candidates." Higher Education **15**(5): 507-522.

Zima, T. (2008). Řád přijímacího řízení 1. lékařské fakulty Univerzity Karlovy v Praze.
Praha, Univerzita Karlova v Praze, 1. lékařská fakulta.