O Teste de Progresso como Indicador para Melhorias em Curso de Graduação em Medicina

Progress Testing as an Indicator for Improvements in a Medical School

Maria Inês da Rosa[I]
Camila Carminati Isoppo[I]
Helen Dominik Cattaneo[I]
Kristian Madeira[I]
Fernando Adami[I]
Olavo Franco Ferreira Filho[II]

PALAVRAS-CHAVE

– Escola Médica.
– Educação.
– Universidade.
– Indicador.
– Qualidade.

RESUMO

Introdução: *O Núcleo de Apoio Pedagógico Interinstitucional Sul II (Napisul II) foi estabelecido em agosto de 2010 com o apoio da Associação Brasileira de Educação Médica (Abem), tendo o propósito específico de formulação, implementação e análise do teste de progresso.* **Objetivo:** *Verificar se os resultados do teste de progresso podem ser usados como indicador para melhorar a qualidade do curso de Medicina.* **Métodos:** *Foram realizados três estudos transversais institucionais durante os três anos da aplicação deste teste no curso de Medicina da Universidade do Extremo Sul Catarinense (Unesc), localizada na Região Sul do Brasil. Todos os participantes do estudo eram estudantes de graduação em Medicina na Unesc e haviam feito o teste de progresso em 2011, 2012 e 2013. A análise estatística foi realizada com um nível de confiança de 95%.* **Resultados:** *A adesão média ao longo dos três anos variou de 91,8% a 100%. Em 2011, o curso de Medicina da Unesc obteve classificação igual ou superior à média dos oito cursos que compõem o Napisul II até a oitava fase e ficou abaixo da média na quinta e sexta fases. Em 2012, a Unesc novamente foi classificada com desempenho mediano até a sétima fase. Na oitava fase, a classificação da Unesc foi significativamente maior do que a média global, e assim, na quinta fase, o curso não diferiu da média do grupo. No entanto, na sexta fase, a classificação do curso foi significativamente inferior à média. Em 2013, a Unesc foi novamente classificada com desempenho mediano até a oitava fase, e nas últimas três fases o curso foi classificado acima da média.* **Conclusão:** *O teste de progresso é um excelente indicador para os gestores, pois pode ser usado para desenvolver intervenções para melhorar a qualidade dos cursos. Após a aplicação do primeiro teste foram realizadas mudanças no curso de Medicina da universidade. Testes posteriores demonstraram a eficácia dessas mudanças.*

Recebido em: 12/09/2016

Aprovado em: 26/09/2016

[I] *Universidade do Extremo Sul Catarinense, Criciúma, SC, Brasil.*
[II] *Universidade Estadual de Londrina, Londrina, PR, Brasil.*

DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

KEYWORDS

– Medical Education.

– Education.

– Universities.

– Indicators.

– Quality.

ABSTRACT

Introduction: *The Núcleo de Apoio Pedagógico Interinstitucional Sul II (Napisul II) was established in August 2010 with the support of the Brazilian Association of Medical Education (Abem) with the specific purpose of formulating, implementing and analyzing the progress test.* **Objective:** *To verify whether the results of the progress test can be used as an indicator to improve the degree course in medicine.* **Methods:** *We performed three cross-sectional studies of institutional households during the three years of the application of the test in the medical school at the University of Extremo Sul Catarinense (Unesc), South Brazil. All participants in the study were undergraduate students in medicine at the Unesc who had taken the progress test in 2011, 2012, and 2013. Statistical analysis was performed with a 95% confidence level.* **Results:** *Mean adherence over the three years ranged from 91.8% to 100%. In 2011, the medical program at the Unesc ranked equal to or above the average for the eight Napisul II schools until the eighth stage and ranked below average in the fifth and sixth years. In 2012, the Unesc again was ranked in the average range until the seventh stage. In the eighth stage, the Unesc's ranking was significantly higher than the overall average, and thus, in the fifth year, the school was on par with the group average. However, in the sixth year, the school's ranking was significantly below average. In 2013, Unesc was again ranked in the middle of the group until the eighth phase, and in the last three phases, the school was ranked above average.* **Conclusion:** *The progress test is an excellent indicator for managers, as it can be used to develop interventions to improve the quality of the courses. After the first test was administered and changes in the school's courses were implemented, subsequent tests demonstrated the effectiveness of the changes.*

INTRODUCTION

Assessment of knowledge remains a major problem in medical education. Progress testing, quite popular in the Netherlands and other countries for several years, only recently has been valued as an evaluation tool in Brazilian educational institutions. Progress testing has been used by medical schools that have implemented curriculum changes, by some graduate programs and by some individual courses[1-6].

This technique was first applied in the early 1970s at the University of Missouri – Kansas City School of Medicine and at the University of Limburg, the Netherlands[7].

Progress testing is a longitudinal assessment process in which medical students participate during their undergraduate years. It is a very valuable instrument for monitoring the students' advancement[8].

The progress test consists of questions that address the medical knowledge considered essential for the last year of the medical program. The content of the test items is based on national curriculum guidelines for medical schools instituted by the Ministry of Education and Culture. The extant literature on medical education suggests that there are different ways of creating and implementing Progress testing[1-6,9].

A progress testing was applied twice a year to all students from 2001 to 2004 who were enrolled at the University of São Paulo School of Medicine. The results suggested a progressive cognitive gain from first to sixth year in all eight tests, that is, there was an improvement within every year < .0001). The improvements were better for basic sciences (taught during the first 2 years), clinical sciences ($P < .0001$), and clerkship or rotation ($P < .0001$). There was no difference in the test performance between men and women[10].

The first interagency nucleus for medical education in Brazil includes seven schools from São Paulo, one from Paraná and one from Santa Catarina. This group of schools is in its tenth edition of the progress test. The second interagency group, the Núcleo de Apoio Pedagógico Interinstitucional Sul II, which is composed of six schools from Santa Catarina and three from Paraná, is in its fourth edition of the progress test. With the support of the Brazilian Association of Medical Education (Abem), another nucleus of schools has been created in other states of the country. In both of the aforementioned groups, the progress test is composed of 120 multiple choice questions, each with four options, that assess student knowledge in seven principal areas of medicine (gynecology and obstetrics, surgery, internal medicine, pediatrics, public health, basic sciences, and bioethics). Every participating institution has a general coordinator and teachers for each principal area of knowledge who prepare questions for the progress test that adhere to a specific format, afford pedagogical rigor and include a matrix of references, presenting the justification for each question and a bibliography that is available to students after the administration of the test. All coordinators and their teams meet on a pre-specified date for a review of the final test. This test is applied during the first week of October in

DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

all the Núcleo de Apoio Pedagógico Interinstitucional Sul II institutions, simultaneously.

The medical school of the University of Extremo Sul Catarinense, one of the institutions in the Núcleo de Apoio Pedagógico Interinstitucional Sul II, participated in all editions of the progress test that were administered between 2011 and 2014. This study aimed to verify that the an analysis of the results of the 2011, 2012 and 2013 editions of the progress test can be used as indicators to improve the graduation rate from University of Extremo Sul Catarinense's medical school.

METHODS

Three cross-sectional observational studies of institutional aggregates were performed, and all student results were analyzed. All students who participated were invited to participate in the test voluntarily from the first to the twelfth stages of the course for the period 2011 to 2013. The scores of those students who answered the test items inappropriately were excluded, for example, those who marked a unique alternative for all questions or more than 50% of the test without responding. The dependent variable was the degree of knowledge in medicine, and the independent variables were the seven areas of medicine and year of graduation. In our group there were two PBL schools and six non-PBL schools.

The items were developed by professors from different subject and sent to a committee of professors from various areas (basic sciences, surgery, medical clinic, obstetrics and gynecology, pediatrics and public health), who chose the items to be applied based on the curriculum determined by the Ministry of Education and Culture for students during undergraduate medical school course. The test consisted of 20 items in each area. The degree of difficulty was assumed by the Commission that prepared the test. The degree of difficulty, discriminative capacity and Cronbach's alpha were calculated retrospectively, using only the responses of the students of the sixth year.

The questions degree of difficulty was calculated by the number of correct answers by students: when less than16% of students answered correctly, they were considered very difficult; between16-40% were considered difficult, between 41-60% were considered medium, between 61-85% were considered easy and more than 85% were classified as very easy. The criteria used to define the questions degree of discrimination were the number of correct answers: < 10% = weak, 10-20% = poor, 21-40% good, 41-60% = very good and >60% = excellent.

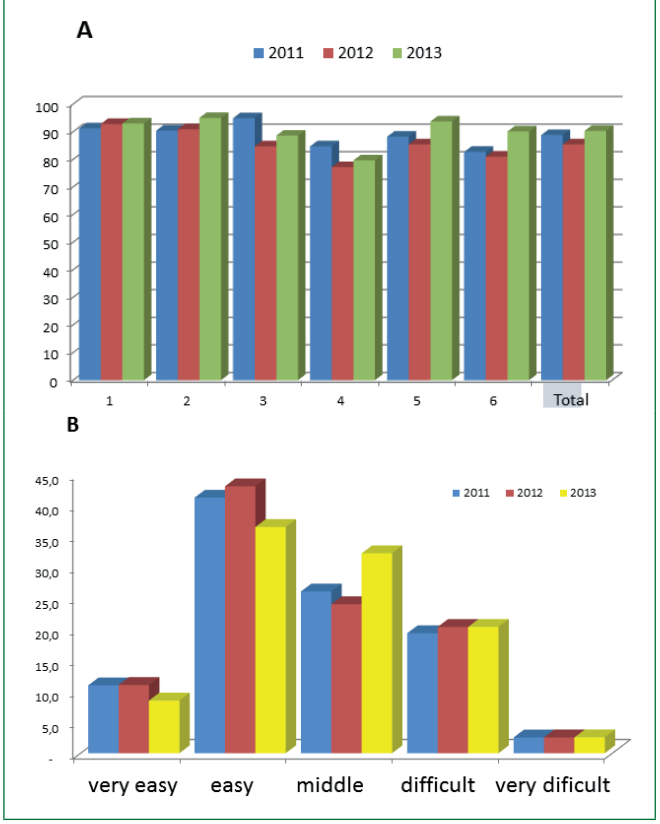
The scores presented are the averages of the percentage of correct answers for each student in the test and in each area of knowledge, and the predictor variable is the stage of the course. There were no significant differences according to gender. The improvement in knowledge was measured according to the analysis of the percentage of correct answers of each year of the course. To compare the mean scores, we used the Mann-Whitney-Wilcoxon test, with a 0.05 level of significance.

The research was approved by the Ethics Committee for Research on Humans of the University of Extremo Sul Catarinense, number 625.470/2014.

RESULTS

The mean participation of students on the progress test was above 90%. Figure 1 shows a predominance of easy questions and means in three years of test application and Figure 2 shows that the degree of discrimination of the questions was good.

Figure 1
A: student presence: means in three years of test application B: degree of discrimination of the questions



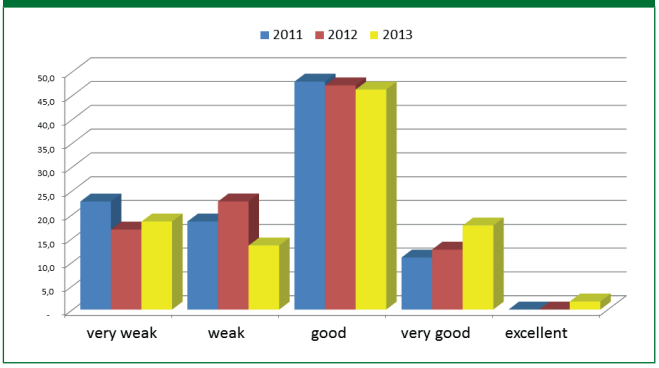
The reliability of tests, internal consistency (Cronbach's α) was above 80% for all years. With respect to first test in 2011, the medical school of University of Extremo Sul Catarinense had 419 (60.2% female) matriculated students, and 98.3% were administered the test.

DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

Table 1
Comparison of the overall performance of Unesc versus Napisul II

	2011 (n) mean ± SD		*p*-value	2012 (n) mean ± SD		*p*-value	2013 (n) mean ± SD		*p*-value
	Unesc	Napisul II		Unesc	Napisul II		Unesc	Napisul II	
1st stage	(34) 39.8 ± 5.3	(113) 40.2 ± 6.6	0.747	(49) 38.6 ± 5.5	(132) 39.3 ± 5.4	0.442	(47) 37.8 ± 5.0	(116) 38.7 ± 5.5	0.333
2nd stage	(36) 43.1 ± 6.6	(324) 43.7 ± 6.7	0.610	(32) 42.0 ± 5.2	(308) 42.7 ± 5.6	0.499	(57) 39.8 ± 5.0	(387) 41.2 ± 5.4	0.066
3rd stage	(40) 44.5 ± 6.9	(107) 44.2 ± 6.3	0.803	(32) 42.4 ± 7.3	(108) 43.0 ± 7.3	0.684	(50) 41.4 ± 5.5	(137) 42.8 ± 6.0	0.151
4th stage	(33) 48.4 ± 6.7	(315) 49.6 ± 6.9	0.341	(35) 46.3 ± 6.8	(319) 47.8 ± 7.4	0.252	(47) 43.7 ± 6.7	(319) 45.9 ± 5.9	0.020
5th stage	(36) 57.7 ± 7.4	(106) 54.7 ± 8.1	0.052	(41) 53.5 ± 8.1	(94) 52.5 ± 8.4	0.521	(47) 49.9 ± 6.4	(88) 49.8 ± 7.8	0.940
6th stage	(34) 58.4 ± 7.9	(326) 57.1 ± 8.8	0.409	(29) 58.4 ± 9.3	(304) 56.1 ± 8.9	0.186	(46) 51.4 ± 6.7	(318) 51.6 ± 8.0	0.872
7th stage	(31) 57.7 ± 7.6	(80) 58.4 ± 9.4	0.712	(35) 61.1 ± 11.4	(78) 60.5 ± 10.0	0.778	(25) 52.7 ± 11.1	(90) 53.5 ± 9.4	0.718
8th stage	(35) 63.3 ± 8.9	(301) 61.4 ± 8.6	0.219	(32) 70.0 ± 9.1	(285) 63.7 ± 11.7	0.003	(43) 57.7 ± 10.9	(282) 57.2 ± 9.6	0.755
9th stage	(37) 57.3 ± 6.8	(97) 61.2 ± 9.7	0.027	(31) 64.6 ± 7.8	(94) 63.3 ± 10.2	0.517	(50) 61.1 ± 7.6	(108) 61.4 ± 9.1	0.840
10th stage	(36) 64.1 ± 7.8	(320) 67.3 ± 9.2	0.046	(35) 72.6 ± 11.3	(281) 68.6 ± 9.5	0.022	(42) 69.3 ± 11.8	(310) 64.1 ± 11.0	0.005
11th stage	(33) 67.0 ± 9.7	(74) 68.1 ± 8.5	0.555	(38) 62.9 ± 9.4	(85) 69.8 ± 11.7	0.002	(32) 64.9 ± 10.5	(93) 62.8 ± 11.3	0.358
12th stage	(27) 65.9 ± 9.2	(276) 71.3 ± 11.8	0.022	(34) 66.1 ± 10.3	(291) 72.5 ± 11.7	0.002	(47) 74.7 ± 14.4	(300) 69.8 ± 13.0	0.018

Figure 2
Discriminative capacity of questions of test items



Compared to the overall average of that schools that made up the Núcleo de Apoio Pedagógico Interinstitucional Sul II, University of Extremo Sul Catarinense recorded a significantly lower average performance in the ninth (p = 0.027), tenth (p = 0.046) and twelfth (p = 0.022) stages and demonstrated no statistically significant difference in the other stages (p >0.05) (Table 1).

Regarding the performance by area in 2011, students in the eighth stage of the University of Extremo Sul Catarinense medical program demonstrated an average number of right answers that was significantly higher than that obtained by other institutions in the Núcleo de Apoio Pedagógico Interinstitucional Sul II in bioethics (p = 0.011), while in the other stages, there were no significant differences (Table 2).

In basic sciences, there was no significant difference between the University of Extremo Sul Catarinense medical students (group 1) and those in the other eight the Núcleo de Apoio Pedagógico Interinstitucional Sul II schools (group 2) (Table 2). Regarding surgery, the group 1recorded a significantly higher than average performance in the fourth (p = 0.001) and fifth (p = 0.038) stages and a lower than average performance in the ninth (0.002) and twelfth (p = 0.001) stages (Table 3) compared to group 2. With respect to internal medicine, in the eighth stage, the University of Extremo Sul Catarinense group performed better than the Núcleo de Apoio Pedagógico Interinstitucional Sul II students (p < 0.001) (Table 3).

The medical students from the University of Extremo Sul Catarinense demonstrated results significantly higher in the fifth (p = 0.001) and sixth (p < 0.001) stages in the area of gynecology and obstetrics but significantly lower (p = 0.014) in the fourth stage (Table 4). On the other hand, there were no statistically significant differences between the two groups of students in the field of pediatrics (Table 5).

In the area of public health, the University of Extremo Sul Catarinense students' performance was significantly lower in the fourth (p = 0.042), fifth (p = 0.024), sixth (p = 0.005), ninth (p = 0.012), tenth (p = 0.001) and twelfth (p = 0.013) stages. In the other stages, there were no significant differences (Table 5).

In 2012, the year of the second progress test, the medical school of University of Extremo Sul Catarinense had 434 (60% female) matriculated students, 97.4% of whom were administered the progress test.

Compared to the overall average of the Núcleo de Apoio Pedagógico Interinstitucional Sul II schools, the University of Extremo Sul Catarinense recorded a significantly higher than average performance in the eighth (p = 0.003) and tenth (p = 0.022) stages, and results below the average performance of

DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

Table 2
Performance comparison of Unesc versus Napisul II by area: Bioethics and Basic Sciencies

	2011 (n) mean ± SD		*p*-value	2012 (n)mean± SD		*p*-value	2013 (n) mean ± SD		*p*-value
	Unesc	Napisul II		Unesc	Napisul II		Unesc	Napisul II	
Bioethics									
1st stage	(34) 2.7 ± 1.3	(113) 2.9 ± 1.3	0.433	(49) 6.0 ± 1.1	(132) 4.0 ± 1.2	<0.001	(47) 2.8 ± 1.1	(116) 3.0 ± 1.0	0.263
2nd stage	(36) 3.2 ± 1.2	(324) 3.1 ± 1.1	0.608	(32) 4.0 ± 1.2	(308) 4.1 ± 1.0	0.598	(57) 2.8 ± 0.8	(387) 2.8 ± 0.9	0.999
3rd stage	(40) 2.8 ± 1.0	(107) 2.9 ± 1.1	0.616	(32) 3.4 ± 1.2	(108) 4.1 ± 1.3	0.007	(50) 2.7 ± 0.9	(137) 3.0 ± 1.0	0.064
4th stage	(33) 3.2 ± 1.2	(315) 3.5 ± 1.2	0.173	(35) 3.9 ± 1.1	(319) 4.5 ± 1.1	0.002	(47) 2.9 ± 1.3	(319) 3.2 ± 1.1	0.089
5th stage	(36) 3.8 ± 1.0	(106) 3.7 ± 1.1	0.631	(41) 3.7 ± 1.2	(94) 4.4 ± 1.2	0.002	(47) 3.1 ± 0.9	(88) 3.2 ± 1.0	0.568
6th stage	(34) 4.0 ± 1.2	(326) 3.9 ± 1.1	0.617	(29) 4.8 ± 0.8	(304) 4.9 ± 1.0	0.602	(46) 3.1 ± 0.9	(318) 3.5 ± 1.1	0.019
7th stage	(31) 3.3 ± 1.2	(80) 3.8 ± 1.2	0.051	(35) 4.9 ± 1.0	(78) 4.9 ± 0.9	0.999	(25) 3.7 ± 1.2	(90) 3.8 ± 1.1	0.694
8th stage	(35) 4.4 ± 1.1	(301) 3.9 ± 1.1	0.011	(32) 5.4 ± 0.8	(285) 5.0 ± 1.0	0.030	(43) 4.2 ± 1.3	(282) 3.8 ± 1.2	0.045
9th stage	(37) 3.6 ± 1.2	(97) 3.7 ± 1.2	0.667	(31) 5.3 ± 0.7	(94) 5.0 ± 0.9	0.093	(50) 4.2 ± 1.2	(108) 4.4± 1.1	0.303
10th stage	(36) 3.9 ± 1.2	(320) 4.0 ± 1.1	0.609	(35) 5.5 ± 0.8	(281) 5.2 ± 0.9	0.061	(42) 4.4 ± 1.2	(310) 3.8 ± 1.2	0.003
11th stage	(33) 4.2 ± 1.1	(74) 4.1 ± 1.0	0.644	(38) 5.0 ± 1.1	(85) 5.2 ± 1.0	0.322	(32) 4.5 ± 0.8	(93) 3.9 ± 1.1	0.005
12th stage	(27) 3.9 ± 1.6	(276) 4.2 ± 1.3	0.264	(34) 5.1 ± 0.9	(291) 5.2 ± 0.8	0.497	(47) 4.8 ± 1.1	(300) 4.1 ± 1.2	<0.001
Basic Sciences									
1st stage	(34) 5.2 ± 1.9	(113) 5.2 ± 1.8	0.999	(49) 4.8 ± 2.0	(132) 5.0 ± 1.2	0.413	(47) 6.2 ± 2.1	(116) 6.3 ± 2.1	0.783
2nd stage	(36) 5.6 ± 1.7	(324) 6.1 ± 1.9	0.131	(32) 7.1 ± 1.9	(308) 6.1 ± 2.2	0.014	(57) 6.5 ± 2.2	(387) 7.5 ± 2.3	0.002
3rd stage	(40) 6.8 ± 1.8	(107) 6.4 ± 2.0	0.270	(32) 6.9 ± 2.5	(108) 7.0 ± 2.3	0.833	(50) 7.1 ± 1.9	(137) 7.6 ± 2.1	0.141
4th stage	(33) 6.7 ± 2.4	(315) 7.1 ± 2.2	0.325	(35) 7.1 ± 1.9	(319) 7.9 ± 2.3	0.048	(47) 7.5 ± 2.6	(319) 8.3 ± 2.4	0.036
5th stage	(36) 8.6 ± 2.3	(106) 7.9 ± 2.1	0.094	(41) 8.6 ± 1.7	(94) 8.6 ± 2.4	0.999	(47) 8.8 ± 2.0	(88) 8.6 ± 2.3	0.616
6th stage	(34) 7.5 ± 2.3	(326) 7.8 ± 2.2	0.452	(29) 8.9 ± 2.0	(304) 8.4 ± 2.4	0.278	(46) 8.4 ± 2.6	(318) 8.8 ± 2.4	0.297
7th stage	(31) 8.0 ± 2.1	(80) 7.7 ± 2.3	0.529	(35) 9.1 ± 2.1	(78) 9.1 ± 2.2	0.999	(25) 8.8 ± 2.0	(90) 9.1 ± 2.3	0.555
8th stage	(35) 9.4 ± 2.1	(301) 8.8 ± 2.3	0.142	(32) 10.5 ± 2.8	(285) 9.4 ± 2.7	0.030	(43) 9.3 ± 2.6	(282) 9.5 ± 2.5	0.627
9th stage	(37) 7.7 ± 1.9	(97) 8.1 ± 1.8	0.259	(31) 8.2 ± 2.5	(94) 8.2 ± 2.5	0.999	(50) 9.9 ± 2.3	(108) 10.2 ± 2.3	0.447
10th stage	(36) 8.2 ± 2.7	(320) 8.7 ± 2.3	0.225	(35) 8.7 ± 2.5	(281) 9.0 ± 2.4	0.488	(42) 10.4 ± 2.4	(310) 9.9 ± 2.5	0.223
11th stage	(33) 8.4 ± 2.5	(74) 8.0 ± 2.2	0.407	(38) 7.7 ± 2.2	(85) 8.8 ± 2.8	0.034	(32) 9.5 ± 2.7	(93) 9.4 ± 2.6	0.853
12th stage	(27) 8.7 ± 2.0	(276) 9.1 ± 2.3	0.384	(34) 8.9 ± 2.0	(291) 9.2 ± 2.4	0.484	(47) 10.4 ± 2.7	(300) 10.4 ± 2.6	0.999

Table 3
Performance comparison of Unesc versus Napisul II by area: Surgery and Intern Medicine

	2011 (n) mean ± SD		*p*-value	2012 (n) mean ± SD		*p*-value	2013 (n) mean ± SD		*p*-value
	Unesc	Napisul II		Unesc	Napisul II		Unesc	Napisul Ii	
Surgery									
1st stage	(34) 5.0 ± 1.5	(113) 5.3 ± 1.9	0.400	(49) 6.5 ± 2.0	(132) 6.1 ± 1.9	0.216	(47) 5.6 ± 1.8	(116) 5.9 ± 2.0	0.374
2nd stage	(36) 5.4 ± 1.8	(324) 5.6 ± 2.0	0.566	(32) 5.8 ± 2.0	(308) 6.7 ± 1.9	0.012	(57) 5.8 ± 1.9	(387) 6.0 ± 1.9	0.459
3rd stage	(40) 5.2 ± 2.3	(107) 5.5 ± 2.2	0.469	(32) 6.2 ± 2.2	(108) 6.0 ± 2.0	0.628	(50) 6.0 ± 1.9	(137) 6.3 ± 1.8	0.322
4th stage	(33) 7.7 ± 2.5	(315) 6.4 ± 2.1	0.001	(35) 7.3 ± 2.0	(319)6.8 ± 2.2	0.199	(47) 6.7 ± 2.0	(319) 6.6 ± 1.9	0.738
5th stage	(36) 7.9 ± 2.7	(106) 6.9 ± 2.4	0.038	(41) 8.3 ± 2.6	(94)7.6 ± 2.5	0.142	(47) 7.6 ± 2.2	(88)7.6 ± 2.0	0.999
6th stage	(34) 7.6 ± 2.1	(326) 7.4 ± 2.3	0.627	(29) 8.9 ± 2.2	(304) 8.4 ± 2.4	0.281	(46) 7.4 ± 2.0	(318) 7.9 ± 2.2	0.146
7th stage	(31) 7.6 ± 1.8	(80)7.8 ± 2.1	0.641	(35) 10.0 ± 3.0	(78) 9.6 ± 2.8	0.494	(25) 7.4 ± 2.8	(90)7.7 ± 2.3	0.584
8th stage	(35) 7.9 ± 2.7	(301) 8.2 ± 2.6	0.520	(32) 11.6 ± 2.3	(285) 9.9 ± 2.8	0.001	(43) 8.1 ± 2.5	(282) 8.9 ± 2.5	0.051
9th stage	(37) 6.5 ± 2.0	(97)8.0 ± 2.6	0.002	(31) 10.3 ± 1.9	(94) 9.4 ± 2.5	0.069	(50) 8.1 ± 2.4	(108) 8.5 ± 2.4	0.331
10th stage	(36) 8.1 ± 2.2	(320) 8.8 ± 2.4	0.095	(35) 11.7 ± 2.4	(281)10.5 ± 2.5	0.008	(42) 8.7 ± 2.5	(310) 9.7 ± 2.4	0.012
11th stage	(33) 8.2 ± 2.2	(74) 8.7 ± 2.1	0.265	(38) 9.7 ± 2.8	(85) 11.5 ± 3.0	0.002	(32) 8.5 ± 1.9	(93)9.2 ± 2.1	0.098
12th stage	(27) 7.9 ± 2.5	(276) 9.7 ± 2.7	0.001	(34)10.8 ± 2.7	(291) 11.7 ± 2.5	0.050	(47) 11.5 ± 2.5	(300) 10.8 ± 2.4	0.065

DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

Table 3
Performance comparison of Unesc versus Napisul II by area: Surgery and Intern Medicine

	2011 (n) mean ± SD		*p*-value	2012 (n) mean ± SD		*p*-value	2013 (n) mean ± SD		*p*-value
	Unesc	Napisul II		Unesc	Napisul II		Unesc	Napisul Ii	
Intern Medicine									
1st stage	(34) 5.7 ± 2.1	(113) 5.7 ± 2.0	0.999	(49) 5.5 ± 1.9	(132) 5.7 ± 2.1	0.560	(47) 4.2 ± 1.8	(116) 4.3 ± 1.9	0.758
2nd stage	(36) 6.3 ± 1.6	(324) 6.1 ± 2.0	0.563	(32) 5.6 ± 1.4	(308) 6.2 ± 1.9	0.083	(57) 4.8 ± 2.0	(387) 4.4 ± 1.9	0.141
3rd stage	(40) 6.8 ± 1.9	(107) 6.7 ± 1.8	0.768	(32) 6.5 ± 2.2	(108) 6.4 ± 2.0	0.809	(50) 4.8 ± 1.8	(137) 4.8 ± 1.8	0.999
4th stage	(33) 7.1 ± 1.8	(315) 6.9 ± 2.1	0.599	(35) 6.9 ± 1.8	(319) 6.8 ± 1.9	0.767	(47) 4.5 ± 2.0	(319) 4.4 ± 1.8	0.726
5th stage	(36) 7.3 ± 2.0	(106) 7.4 ± 2.1	0.803	(41) 7.2 ± 2.3	(94) 7.4 ± 2.2	0.633	(47) 5.0 ± 2.5	(88) 5.0 ± 2.0	0.999
6th stage	(34) 8.3 ± 2.0	(326) 8.2 ± 2.2	0.799	(29) 7.9 ± 2.0	(304) 7.9 ± 2.2	0.999	(46) 5.9 ± 1.8	(318) 5.9 ± 2.2	0.999
7th stage	(31) 8.5 ± 2.1	(80) 8.3 ± 2.1	0.653	(35) 8.5 ± 2.1	(78) 8.6 ± 2.1	0.815	(25) 6.1 ± 2.5	(90) 6.2 ± 2.3	0.851
8th stage	(35) 10.0 ± 2.0	(301) 8.6 ± 2.2	<0.001	(32) 10.9 ± 2.4	(285) 9.5 ± 2.3	0.001	(43) 8.2 ± 2.6	(282) 7.4 ± 2.5	0.053
9th stage	(37) 8.2 ± 1.7	(97) 8.6 ± 2.2	0.320	(31) 9.4 ± 2.4	(94) 9.3 ± 2.4	0.841	(50) 7.7 ± 2.2	(108) 7.1 ± 2.3	0.124
10th stage	(36) 9.8 ± 1.7	(320) 9.7 ± 2.2	0.792	(35) 11.1 ± 2.5	(281) 10.2 ± 2.4	0.038	(42) 8.5 ± 2.9	(310) 8.7 ± 2.7	0.656
11th stage	(33) 10.4 ± 2.1	(74) 10.1 ± 2.0	0.482	(38) 9.5 ± 2.4	(85) 10.4 ± 2.6	0.072	(32) 8.9 ± 3.2	(93) 8.2 ± 3.2	0.288
12th stage	(27)10.0 ± 2.8	(276) 10.4 ± 2.7	0.465	(34) 9.7 ± 2.5	(291) 11.1 ± 2.2	0.001	(47) 11.2 ± 3.6	(300) 9.8 ± 3.1	0.005

Table 4
Performance comparison of Unesc versus Napisul II by area: Gynecology and Obstetrics

	2011 (n) mean ± SD		*p*-value	2012 (n) mean ± SD		*p*-value	2013 (n) mean ± SD		*p*-value
	Unesc	Napisul II		Unesc	Napisul II		Unesc	Napisul II	
1st stage	(34) 6.1 ± 1.7	(113)6.5 ± 2.0	0.293	(49) 5.4 ± 2.0	(132) 5.6 ± 1.8	0.520	(47) 5.2 ± 2.0	(116) 5.3 ± 1.7	0.747
2nd stage	(36) 6.3 ± 2.2	(324) 6.6 ± 2.1	0.419	(32) 5.8 ± 1.3	(308) 6.0 ± 1.9	0.562	(57) 5.6 ± 2.1	(387) 5.8 ± 1.8	0.444
3rd stage	(40) 6.6 ± 2.4	(107) 6.5 ± 2.2	0.811	(32) 6.3 ± 1.9	(108) 5.8 ± 2.0	0.211	(50) 6.1 ± 1.8	(137) 5.9 ± 1.9	0.519
4th stage	(33) 6.7 ± 2.3	(315) 7.7 ± 2.2	0.014	(35) 6.1 ± 2.0	(319) 7.0 ± 2.3	0.027	(47) 6.2 ± 1.7	(319) 6.9 ± 1.9	0.017
5th stage	(36)11.1 ± 2.4	(106) 9.4 ± 2.7	0.001	(41) 9.4 ± 2.0	(94) 8.2 ± 2.6	0.009	(47) 8.4 ± 1.9	(88) 8.1 ± 2.3	0.446
6th stage	(34)10.6 ± 2.0	(326) 9.1 ± 2.3	< 0.001	(29) 10.5 ± 2.5	(304) 8.9 ± 2.6	0.002	(46) 9.9 ± 1.9	(318) 8.1 ± 2.2	< 0.001
7th stage	(31)9.9 ± 1.5	(80) 10.3 ± 2.1	0.335	(35) 10.0 ± 3.0	(78) 10.0 ± 2.5	0.999	(25) 8.8 ± 2.7	(90) 8.7 ± 2.5	0.862
8th stage	(35) 9.9 ± 2.0	(301) 10.2 ± 2.3	0.460	(32) 11.0 ± 2.1	(285) 10.5 ± 2.5	0.277	(43) 9.6 ± 2.8	(282) 9.2 ± 2.3	0.304
9th stage	(37) 11.6 ± 2.2	(97)11.9 ± 2.4	0.509	(31) 12.5 ± 2.7	(94) 11.6 ± 2.6	0.100	(50) 12.0 ± 2.3	(108) 11.1 ± 2.5	0.033
10th stage	(36) 12.2 ± 2.0	(320) 12.3 ± 2.3	0.802	(35) 12.9 ± 2.0	(281) 12.3 ± 2.3	0.141	(42) 13.5 ± 2.7	(310) 11.0 ± 2.7	< 0.001
11th stage	(33) 11.5 ± 2.3	(74) 12.3 ± 2.2	0.090	(38) 11.4 ± 2.5	(85) 12.3 ± 2.5	0.068	(32) 12.2 ± 2.7	(93) 11.7 ± 2.8	0.381
12th stage	(27) 11.7 ± 2.7	(276) 12.8 ± 3.0	0.068	(34) 11.4 ± 2.0	(291) 12.7 ± 2.4	0.003	(47) 13.1 ± 3.1	(300) 12.2 ± 2.9	0.051

DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

Table 5
Performance comparison of Unesc versus Napisul II by area: Pediatrics and Public Health

	2011 (n) mean ±SD		*p*-value	2012 (n) mean ±SD		*p*-value	2013 (n)mean± SD		*p*-value
	Unesc	Napisul II		Unesc	Napisul II		Unesc	Napisul II	
Pediatrics									
1st stage	(34) 6.2 ± 1.4	(113) 6.1 ± 1.9	0.777	(49) 5.6 ± 1.9	(132) 5.8 ± 2.0	0.545	(47) 6.2 ± 2.0	(116) 6.8 ± 2.0	0.085
2nd stage	(36) 6.9 ± 1.8	(324) 6.6 ± 2.0	0.389	(32) 6.4 ± 1.9	(308) 6.3 ± 2.0	0.787	(57) 7.5 ± 2.3	(387) 7.2 ± 2.1	0.321
3rd stage	(40) 7.2 ± 1.8	(107) 6.6 ± 2.0	0.099	(32) 6.5 ± 2.4	(108) 6.4 ± 2.3	0.831	(50) 7.5 ± 2.0	(137) 7.6 ± 2.1	0.771
4th stage	(33) 7.6 ± 1.8	(315) 7.6 ± 2.3	0.999	(35) 8.3 ± 2.2	(319) 7.4 ± 2.4	0.034	(47) 8.4 ± 2.1	(319) 8.1 ± 2.3	0.399
5th stage	(36)8.9 ± 2.3	(106) 8.4 ± 2.1	0.230	(41) 9.2 ± 2.6	(94) 8.3 ± 2.8	0.082	(47) 9.5 ± 1.8	(88) 9.1 ± 2.2	0.287
6th stage	(34)10.0 ± 2.2	(326) 9.3 ± 2.3	0.091	(29) 10.1 ± 2.2	(304) 9.1 ± 2.7	0.054	(46) 9.2 ± 2.2	(318) 8.8 ± 2.2	0.250
7th stage	(31)9.9 ± 2.1	(80) 9.5 ± 2.5	0.432	(35) 10.9 ± 2.9	(78) 10.4 ± 2.9	0.399	(25) 9.3 ± 2.2	(90) 9.3 ± 2.0	0.999
8th stage	(35) 10.3 ± 2.6	(301) 9.9 ± 2.4	0.356	(32) 12.1 ± 2.0	(285) 10.4 ± 3.0	0.002	(43) 9.9 ± 2.3	(282) 9.7 ± 2.2	0.581
9th stage	(37) 8.9 ± 3.0	(97)10.0 ± 3.0	0.060	(31) 10.5 ± 2.3	(94) 11.1 ± 2.6	0.254	(50) 9.6 ± 2.1	(108) 10.3 ± 2.2	0.061
10th stage	(36) 10.9 ± 2.5	(320) 11.2 ± 2.5	0.495	(35) 12.9 ± 2.4	(281) 11.7 ± 2.6	0.010	(42) 12.6 ± 2.8	(310) 11.0 ± 2.7	< 0.001
11th stage	(33) 12.2 ± 2.5	(74) 11.9 ± 2.2	0.534	(38) 11.5 ± 2.7	(85) 12.3 ± 2.7	0.132	(32) 11.8 ± 2.2	(93) 11.1 ± 2.4	0.149
12th stage	(27) 11.6 ± 1.8	(276) 11.7 ± 2.5	0.840	(34) 11.3 ± 2.6	(291) 12.4 ± 2.8	0.030	(47) 12.4 ± 2.0	(300) 11.7 ± 2.5	0.068
Public Health									
1st stage	(34) 8.9 ± 2.2	(113) 8.5 ± 2.4	0.387	(49) 7.4 ± 1.7	(132) 7.2 ± 1.6	0.464	(47) 7.6 ± 2.1	(116) 7.2 ± 2.0	0.256
2nd stage	(36) 9.4 ± 2.0	(324) 9.7 ± 2.3	0.453	(32) 7.3 ± 1.6	(308) 7.4 ± 1.9	0.774	(57) 6.9 ± 1.7	(387) 7.5 ± 2.1	0.040
3rd stage	(40) 9.2 ± 2.6	(107) 9.7 ± 2.3	0.260	(32) 6.7 ± 1.7	(108) 7.3 ± 1.8	0.096	(50) 7.1 ± 2.2	(137) 7.5 ± 2.0	0.240
4th stage	(33) 9.5 ± 2.5	(315) 10.4 ± 2.4	0.042	(35) 6.6 ± 2.0	(319) 7.4 ± 2.2	0.040	(47) 7.6 ± 2.3	(319) 8.3 ± 2.2	0.044
5th stage	(36) 10.1 ± 1.8	(106) 11.1 ± 2.3	0.024	(41) 7.1 ± 1.7	(94) 7.9 ± 1.9	0.022	(47) 7.4 ± 2.8	(88) 8.2 ± 2.6	0.100
6th stage	(34) 10.4 ± 2.3	(326) 11.6 ± 2.5	0.005	(29) 7.3 ± 1.9	(304) 8.3 ± 2.0	0.010	(46) 7.5 ± 1.5	(318) 8.6 ± 2.3	0.002
7th stage	(31) 10.4 ± 2.5	(80) 11.1 ± 2.5	0.188	(35) 7.7 ± 1.8	(78) 7.9 ± 1.8	0.586	(25) 8.7 ± 2.9	(90) 8.6 ± 2.7	0.872
8th stage	(35) 11.5 ± 2.1	(301) 11.7 ± 2.4	0.650	(32) 8.6 ± 2.2	(285) 8.8 ± 2.3	0.640	(43) 8.3 ± 2.4	(282) 8.7 ± 2.4	0.309
9th stage	(37) 10.8 ± 2.1	(97) 11.9 ± 2.3	0.012	(31) 8.5 ± 1.9	(94) 8.7 ± 2.3	0.663	(50) 9.6 ± 1.9	(108) 9.9 ± 2.5	0.452
10th stage	(36) 11.2 ± 2.0	(320) 12.6 ± 2.3	0.001	(35) 9.8 ± 2.7	(281) 9.8 ± 2.3	0.999	(42) 11.3 ± 2.7	(310) 10.0 ± 2.7	0.004
11th stage	(33) 12.2 ± 2.3	(74) 12.9 ± 2.3	0.149	(38) 8.2 ± 1.8	(85) 9.3 ± 2.4	0.013	(32) 9.5 ± 1.7	(93) 9.4 ± 2.4	0.828
12th stage	(27) 12.1 ± 2.5	(276) 13.4 ± 2.6	0.013	(34) 8.6 ± 2.5	(291) 10.2 ± 2.6	0.001	(47) 11.3 ± 2.7	(300) 10.8 ± 3.0	0.283

the Núcleo de Apoio Pedagógico Interinstitucional Sul II schools in the eleventh (p = 0.002) and twelfth (p = 0.002) stages. There were no significant differences in the other stages (p >0.05) (Table 1).

In relation to performance by area, in bioethics (Table 2), the University of Extremo Sul Catarinense recorded an average number of right answers that was significantly higher than that of the Núcleo de Apoio Pedagógico Interinstitucional Sul II institutions in the first (p < 0.001) and eighth (p = 0.030) stages, while in the third (p = 0.007), fourth (p = 0.002) and fifth (p = 0.002) stages, the University of Extremo Sul Catarinense group demonstrated a significant lower performance than that of the Núcleo de Apoio Pedagógico Interinstitucional Sul II group.

In basic sciences, the University of Extremo Sul Catarinense students recorded significant higher performances in the second (p = 0.014) and eighth (p = 0.014) stages but lower performances in the fourth (p = 0.048) and eleventh (p = 0.034) stages compared to their the Núcleo de Apoio Pedagógico Interinstitucional Sul II counter parts (Table 2).

In the field of surgery, the University of Extremo Sul Catarinense students performed significantly below the Núcleo de Apoio Pedagógico Interinstitucional Sul II students in the second (p = 0.012), eleventh (p = 0.002) and twelfth (p = 0.050) stages but significantly above their counter parts in the eighth (p = 0.001) and tenth (p = 0.008) stages (Table 3).

In internal medicine, the University of Extremo Sul Catarinense group scored significantly higher than the Núcleo de Apoio Pedagógico Interinstitucional Sul II group in the eighth (p = 0.001) and tenth (p = 0.038) stages and scored below them in the twelfth (p = 0.001) stage (Table 3).

In the area of gynecology and obstetrics (Table 4), University of Extremo Sul Catarinense demonstrated significantly higher results in the fifth and sixth stages (p = 0.009 and p = 0.002, respectively) but significantly lower in the fourth (p = 0.027) and twelfth (p = 0.003) stages.

In pediatrics, the difference between the two groups was significantly lower only in the twelfth (p = 0.030) stage, while it was significantly and higher in the fourth, eighth and

DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

tenth stages (p = 0.034, p = 0.002 and p = 0.010, respectively) (Table 5).

In public health, the University of Extremo Sul Catarinense group recorded significantly lower scores than the Núcleo de Apoio Pedagógico Interinstitucional Sul II group in the fourth (p = 0.040), fifth (p = 0.022), sixth (p = 0.010), eleventh (p = 0.013) and twelfth (p = 0.001) stages (Table 5).

In 2013, when the third test was administered, the medical school of University of Extremo Sul Catarinense had 474 (59.4% female) matriculated students, 96.8% participated in Progress testing.

When comparing the overall averages of the University of Extremo Sul Catarinense medical students with those of the Núcleo de Apoio Pedagógico Interinstitucional Sul II medical students (Table 1), the University of Extremo Sul Catarinense group recorded significantly higher scores in the tenth and twelfth stages (p = 0.005 and p = 0.018, respectively) and significantly lower scores in the fourth stage (p = 0.020). The other stages presented no statistically significant differences (p >0.050).

With respect to performance by area, the University of Extremo Sul Catarinense students demonstrated a significantly lower than average performance in bioethics (Table 2) in the sixth stage (p = 0.019) compared to the Núcleo de Apoio Pedagógico Interinstitucional Sul II average but performed better than the Núcleo de Apoio Pedagógico Interinstitucional Sul II average in the eighth, tenth, eleventh and twelfth (p = 0.045, p = 0.003, p = 0.005, p < 0.001, respectively) stages. In basic sciences, the average for the University of Extremo Sul Catarinense group was significant lower in the second (p = 0.002) and fourth (p = 0.036) stages than it was for the Núcleo de Apoio Pedagógico Interinstitucional Sul II group, and there were no significant differences in the other stages.

In surgery (Table 3), the University of Extremo Sul Catarinense average in relation to the Núcleo de Apoio Pedagógico Interinstitucional Sul II average was significantly lower only in the tenth stage (p = 0.012), and no significant differences were noted for the remaining stages. Regarding intern medicine, a statistically significant difference was observed between the two groups in the twelfth stage (p = 0.005), with the University of Extremo Sul Catarinense medical school students recording a higher average score. There were no observed significant differences in the other stages (Table 3).

In gynecology and obstetrics, scores of students at University of Extremo Sul Catarinense were significantly lower than those of students in the Núcleo de Apoio Pedagógico Interinstitucional Sul II schools in the fourth stage (p = 0.017), whereas University of Extremo Sul Catarinense's scores in the sixth (p < 0.001), ninth (p = 0.033) and tenth (p < 0.001) stages were significantly above those of the Núcleo de Apoio Pedagógico Interinstitucional Sul II schools (Table 4).

In pediatrics, there was a significantly higher difference in the average scores in the tenth stage (p < 0.001) in favor of the University of Extremo Sul Catarinense, while the other stages demonstrated no differences compared to the Núcleo de Apoio Pedagógico Interinstitucional Sul II institutions (Table 5).

Finally, with respect to public health (Table 5), the University of Extremo Sul Catarinense recorded a higher average score that the Núcleo de Apoio Pedagógico Interinstitucional Sul II in the tenth (p = 0.004) stage and lower average scores in the second, fourth and sixth stages (p = 0.040, p = 0.044, and p = 0.002, respectively).

DISCUSSION

Our study found that Progress testing proved to be an important indicator for decision makers working to improve the medical school at the University of Extremo Sul Catarinense.

Progress testing is a form of assessment in which the test is administered simultaneously to students at different stages of their training as a way to monitor progress over time. The test is defined by a level of difficulty equivalent to an end of training assessment, and it is at this level that all students, from the first to the twelfth stage, are tested. We use four alternatives for the answers. We did not use "I do not know" option to answers as suggested by some authors[11].

Some limitations of this study must be considered. First, it combines three cross-sectional studies in a temporal analysis, and the degree of difficulty of the tests differs each year. Thus, the growth curves of cognitive knowledge of students when compared year-by-year should be analyzed sparingly. To minimize these biases, a statistical analysis that includes using tests that measure the comparability between different years of tests is used to improve accuracy, for example, the anchorage test[12]. Concerned about the equivalency between and among the tests, the Núcleo de Apoio Pedagógico Interinstitucional Sul II applies tests of equivalence. The results of these tests will be presented in another article. Despite its limitations, the progress test, in addition to providing managers of courses with evidence for curricula decision-making, the test also provides valuable individual feedback, thus enhancing student learning. Some studies demonstrated that repeated testing with feedback appears to result, for the student, in greater retention of long-term knowledge, and thus, they suggested that Progress testing should be considered important not only as an evaluation tool but also as a powerful learning tool[13].

 DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

Although the PT has been structured to evaluate schools using problem-based learning methodology, it showed that it can be used in school with different curriculum. The knowledge acquisition of international mobility medical students among Dutch, German and Italian medical schools was compared using the Mastricht Progress Testing. They concluded that this test could be used to identify corresponding cognitive levels of students on individual domains in different curricula[14,15].

A study performed with students of all six classes of two medical schools in the Netherlands, one PBL and non-PBL, showed no differences in the total test scores. After separating the test in three categories (basic, clinical and social sciences) a few but non-systematic differences were found. To some extent the non-PBL students scored better on the basic sciences while the PBL students had better scores on the social sciences[16].

In a study conducted at the medical school of McMaster University, Canada, the immediate effect of the introduction of Progress testing was a reduction in the failure rate of students from that institution on the national licensing exam from 19% to 4.5%[17].

A study conducted in California assessed the correlation of the number of months in clinical training with clinical knowledge measured by the United States Medical licensing Examination, it showed a highly significant correlation between increasing the months of clinical training and the acquisition and utilization of clinical knowledge[18].

McMaster PT was introduced in 1991 as a required assessment to distinguish it from traditional examinations. After two years in use the TP remained a guide to personal progress[19]. In Brazil, before the formation of the Núcleo de Apoio Pedagógico Interinstitucional Sul II, some institutions had adopted routine Progress testing, and the results of this practice, published in the Brazilian Congress of Medical Education, emphasized the importance of Progress testing in the evaluation of schools and students. Accordingly, the concept of Progress testing created great expectations among both teachers and students as Progress testing began to be implemented in the medical schools of Parana and Santa Catarina. The eight institutions that make up the nucleus joined in the preparation of tests. They formed a cohesive group that was then enabled pedagogically to prepare items, and they were determined to provide both quantitative and qualitative items that would ultimately enhance student performance. Thus, the coordinators of each institution were eager to receive the test results, which were always conducted confidentially and individually and did not allow for any "ranking" among schools. The analyses were performed by an external service contracted by the group.

With results in hand, the school coordinator of the University of Extremo Sul Catarinense first met with his team of teachers and coordinators to discuss, in broad terms, the results and to define strategies for improvement, which were later presented collegiately with student participation.

On the first progress test (2011), the overall average for the medical school of University of Extremo Sul Catarinense, compared to the overall average for the Núcleo de Apoio Pedagógico Interinstitucional Sul II schools, had a significantly lower number of right answers in the area of medical internship (ninth, tenth and twelfth stages), presented no differences in the other stages, and performed below expectations in most stages with respect to in public health. Carefully analyzing the school educational plan, it was concluded that in the stages where problem-based learning (PBL) methodology was incorporated, the first to the eighth stage, the average scores for the University of Extremo Sul Catarinense students was equivalent to the average of the Núcleo de Apoio Pedagógico Interinstitucional Sul II schools, but during the medical internship, the scores for the University of Extremo Sul Catarinense students were below average. With respect to public health, although students had been inserted into the primary healthcare network since the first stage, it was determined that the theoretical-conceptual content was insufficient.

Thus, it was decided to insert weekly activities in the form of dialogued lectures, seminars, discussions and critical analyses of articles and texts of general content in addition to an hour of themes addressing public health as part of the assessment of medical internship. In the other stages, the implementation of seminars about public health by teachers in the basic units of health was encouraged.

On the second edition of the progress test, the University of Extremo Sul Catarinense demonstrated growth in learning by achieving a significantly higher than average performance in the eighth and tenth stages, while it performed in the average range in the others stages, with the exception of the last year. The University of Extremo Sul Catarinense's students did not have a good performance in the progress test in last year of course perhaps because interventions were implemented at the beginning of the internship. Furthermore, the school recorded a lower performance on the progress test in the eleventh and twelfth stages.

On the other hand, in 2013, the overall average performance of the University of Extremo Sul Catarinense in relation to the other eight schools that make up the Núcleo de Apoio Pedagógico Interinstitucional Sul II exhibited a significantly greater difference in the tenth and twelfth stages (final year), while in the ninth and tenth stages, the University of Extremo

 DOI: http://dx.doi.org/10.1590/1981-52712015v41n1RB20160022

Sul Catarinense's performance was on average with the other eight schools.

A Cross-sectional study was conducted to investigate the potential of PT in developing academic management. The performance of medical students were analyzed and applied in 2008 and reapplied unchanged in 2011. The results showed that, in 2008, the PT progression occurred every two years, and in 2011 this progression started in the third year. In the areas of Public Health, Surgery and Basic Areas, no gradual improvements of knowledge was observed from initial years, which deserve reflections on the part of curriculum managers[20].

Therefore, with respect to the University of Extremo Sul Catarinense, the results of the progress test that was administered from 2011 to 2013, when carefully analyzed by the management team, indicate that the test is an important tool for improving the medical school, especially in relation to the internship, which is the area where the test indicated the greatest shortcomings. However, the major difficulty for managers is to compare year-to-year results because, even though the tests are developed by the same team, the tests are not identical, and thus, there remains potential bias in that advancements in student performance as well as declines in student performance are due to the level of difficulty of test items from one test to another. As a future perspective, there is the need to prepare and analyze the TP questions to ensure that the tests maintain a consistent level of difficulty such that the results are determined exclusively by student performance and not by the degree of the test's difficulty.

REFERENCES

1. Aarts R, Steidell K, Manuel BAF, Driessen EW. Progress testing in resource-poor countries: a case from Mozambique. *Medical Teacher* 2010; (32):461–3.
2. Bennett J, Freeman A, Coombs L, Kay L, Ricketts C. Adaptation of medical progress testing to a dental setting. *Medical Teacher* 2010; (32):500–2.
3. Freeman A, Van der Vleuten C, Nouns Z, Ricketts C. Progress testing internationally. *Medical Teacher* 2010; (32):451–5.
4. Nouns Z, Georg W. Progress testing in German speaking countries. *Medical Teacher* 2010; (32):467–70.
5. Schuwirth L, Bosman G, Henning R, Rinkel R, Wenink A. Collaboration on progress testing in medical schools in the Netherlands. *Medical Teacher*2010; (32):476–9.
6. Swanson D, Holtzman K, Butler A, Langer MM, Nelson MV, Chow JW et al. Collaboration across the pond: the multi-school progress testing project. *Medical Teacher* 2010; (32):480–5.
7. Blake JM, Norman GR, Keane DR, Mueller B, Cunnington J, Didyk N. Introducing progress testing in McMaster University's problem-based medical curriculum: psychometric properties and effect on learning. *Academic Medicine* 1996; 71(9):1002-7.
8. Boshuizen HP, Van der Vleuten CPM, Schmidt HG, Machiels-Bongaerts M. Measuring knowledge and clinical reasoning skills in a problem-based learning curriculum. Med Educ. 1997;31(2):115-21.
9. Van der Vleuten CPM, Verwijnen GM, Wijnen HFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Medical Teacher* 1996; (18):103–10.
10. Tomic ER, Martins MA, Lotufo PA, Benseñor IM. Progress testing: evaluation of four years of application in the school of medicine, University of São Paulo. Clinics (Sao Paulo). 2005;60:389-96.
11. Muijtjens AMM, Van Mameren H, Hoogenboom RJI, Evers JLH, Van der Vleuten CPM. The effect of "don't know" option on test scores: Number-right and formula scoring compared. Med Educ. 1999;33(4):267-75.
12. Pasquali L. Psicometria- teoria dos testes na psicologia e na educação, Petrópolis-RJ: *Vozes*, 2003; 261-278.
13. Larsen DP, Butler AC, Roediger HL. Repeated testing improves long-term retention relative to repeated study: a randomized controlled trial. *Medical Education* 2009; (43):1174–81.
14. Albano MG, Cavallo F, Hoogenboom R, Magni F, Majoor G, Manenti F et al. An international comparison of the knowledge levels of medical students: the Maastricht progress test. Med Educ. 1996;30(4):239-45.
15. Schmidt HG. Innovative and conventional curricula compared: what can be said about their effects? In: Nooman ZM, Schmidt HG, Ezzat ES, editors. Innovation in medical education. An evaluation of its present status. New York: Springer Publishing Company; 1990. p. 1-7.
16. Verhoeven BH, Verwijnen GM, Scherphier AJJA, Holdrinet RSG, Oeseburg BM, Bulte JA, et al. An analysis of progress test results of PBL and non-PBl students. In: van der Vleuten CPM, Scherpbier AJJA, Wijnen WHFW, editors. Progress testing: the utility of an assessment concept. Groningen: Bas Verhoeven; 2003:125-36.
17. Norman G, Neville A, Blake J, Mueller B. Assessment steers learning down the right road: impact of progress testing on licensing examination performance. *Medical Teacher* 2010; (32):496–9.
18. Vosti KL, Bloch DA, Jacobs CD. The relationship of clinical knowledge to months of clinical training among medical students. Acad Med. 1997;72(4):1097-102.

19. Blake JM, Norman GR, Smith EK. Report card from McMaster: Student evaluation at a problem-based medical school. Lancet. 1995;345(8954):899-902.
20. Pinheiro OL, Spadella MA, Moreira HM, Ribeiro ZMT, Guimarães APC,, Almeida Filho OM, Hafner, Maria de MLMB. (2015). Teste de Progresso: uma Ferramenta Avaliativa para a Gestão Acadêmica. Rev Brasil Educ Méd 2015; 39: 68-78.

CONTRIBUIÇÃO DOS AUTORES

Maria Inês da Rosa: created the idea that originated the work and developed hypotheses; coordinated the group that carried out the work. Camila Carminati Isoppo: wrote the manuscript; reviewed the literature. Helen Dominik Cattaneo: collected data; wrote the manuscript. Kristian Madeira: analyzed the results statistically; has structured the working method. Fernando Adami: reviewed the literature; presented important suggestions incorporated into the work. Olavo Franco Ferreira Filho: guided the writing of the manuscript; presented important suggestions incorporated into the work.

CONFLITO DE INTERESSES

The authors declare no conflicts of interest

ENDEREÇO PARA CORRESPONDÊNCIA

Maria Inês da Rosa
Rua Cruz e Souza, 510
Pio Correa – Criciúma
CEP 88811-550 – SC
E-mail: mir@unesc.net**;** carminati.camila@hotmail.com; helen_dominik@hotmail.com; kristianmadeira@gmail.com; adamifernando@uol.com.br; olavofranco@gmail.com