

Artículos Científicos

Pruebas frecuentes y parciales como indicadores de los resultados finales de la asignatura Metabolismo Intermediario y su Regulación

Dariel Díaz Arce*
Martha Cecilia Olivera de la Torre**

*Lic. en Bioquímica, Profesor Asistente, Departamento de Bioquímica.
**Lic. en Bioquímica, Profesor Asistente, Departamento de Bioquímica.
Institución: Escuela Latinoamericana de Medicina (ELAM).

Resumen

Objetivos: Determinar el grado de asociación entre los resultados de las pruebas frecuentes (corte evaluativo o C1) y pruebas parciales (prueba intrasemestral o PIS y trabajo de control en clase o TCC) con los del examen final ordinario (EF) de la asignatura Metabolismo Intermediario y su Regulación y evaluar, mediante el cálculo de indicadores básicos del desempeño de las pruebas diagnósticas, el valor de la PIS, TCC y C1 como diagnosticadores de los resultados en el examen ordinario de dicha asignatura en la ELAM.

Métodos: Se realizó un estudio comparativo transversal que involucró los resultados docentes del C1, PIS, TCC y EF de 2 523 estudiantes matriculados en el segundo semestre de la carrera de medicina, de los cursos académicos comprendidos entre septiembre de 2001 y julio de 2006.

Resultados: En todos los cursos se encontró una asociación significativa entre los resultados obtenidos en el EF y los esperados, según las pruebas C1, PIS y TCC. Los indicadores de sensibilidad y valor predictivo negativo fueron mejores en la PIS, no así los de especificidad y valor predictivo positivo que fueron mejores en el TCC.

Conclusiones: Este trabajo muestra que la Prueba Intrasemestral en la ELAM es el mejor predictor de los resultados finales de la asignatura Metabolismo Intermediario y su Regulación.

Palabras clave: Bioquímica, evaluación, aprendizaje, educación superior .

INTRODUCCIÓN

La disciplina Bioquímica ha sido reconocida por varios autores como de difícil comprensión y asimilación por los estudiantes de las Ciencias Médicas (1, 2). En algunas de estas carreras se considera entre las principales materias que provocan pérdida del cupo académico (3, 4). Detectar tempranamente a los estudiantes

con dificultades en esta asignatura y desarrollar una intervención que posibilite el cumplimiento de los objetivos instructivos y educativos de la misma, pudiera garantizar que los alumnos culminen satisfactoriamente el curso (1, 5).

Algunas investigaciones se han encaminado a la búsqueda de buenos indicadores de los resultados docentes en el examen final ordinario (EF) de esta disciplina. De este modo, se han obtenido

asociaciones significativas entre algunas variables académicas, intelectuales y psicológicas, con el éxito o fracaso en la prueba final de Bioquímica, en diferentes carreras de las Ciencias Médicas (6, 7, 8).

En la Escuela Latinoamericana de Medicina (ELAM), la cuantificación de algunas de estas variables es extremadamente difícil y en otros casos imposible por la heterogeneidad de sus estudiantes así como por la gran cantidad de ellos. Por eso, en esta institución ha sido necesaria la búsqueda de indicadores sensibles de resultados negativos en el EF de Bioquímica.

Un trabajo inicial desarrollado por Díaz *et al.* (9), mostró que la Prueba Intrasemestral (PIS) es un buen predictor de resultados adversos en el EF de Biología Celular y Molecular (BCM) en la ELAM. Un estudio similar no ha sido desarrollado aún para la asignatura Metabolismo Intermediario y su Regulación (MIR).

Esta investigación tuvo como objetivos determinar el grado de asociación entre los resultados de las pruebas frecuentes: corte evaluativo (C1) y pruebas parciales (PIS y trabajo de control en clase o TCC) con los del EF de MIR en la ELAM y evaluar mediante el cálculo de indicadores básicos del desempeño de las pruebas diagnósticas, el valor de la PIS, TCC y C1 como diagnosticadores de los resultados en el examen ordinario de MIR.

MATERIALES Y MÉTODOS

Se realizó un estudio comparativo transversal que involucró los resultados docentes del C1, PIS, TCC y EF de 2 523 estudiantes. Este tamaño de muestra se calculó teniendo en cuenta los resultados previos de sensibilidad y especificidad global aportados por Pedrosa *et al.* (datos no publicados) con una precisión absoluta menor del 10% y una prevalencia histórica de suspensos del 30%. Estos estudiantes se seleccionaron por un muestreo por conglomerados (grupos académicos) al azar. La población estudiada correspondió a los alumnos matriculados en el segundo semestre de los cursos académicos comprendidos entre septiembre de 2001 y julio de 2006.

Aquellos estudiantes que carecían de alguna evaluación se les consideró como de mal (2 puntos).

Para la determinación de las medidas de asociación entre las variables C1 vs. EF; PIS vs. EF y TCC vs. EF, se emplearon los coeficientes de correlación de Spearman (r_s) y Kappa (K) de Cohen. Para el cálculo de la r_s y la K, los datos del C1 se convirtieron de una escala de B, R y M a 3, 2 y 1, respectivamente. Buscando una homogeneidad en las posteriores comparaciones, los datos de la PIS, TCC y EF también fueron convertidos a la misma escala que el C1, pero considerando que las notas de 4 ó 5 se corresponden a B, 3 a R y 2 a M.

Para determinar la fortaleza de asociación entre las pruebas estudiadas se empleó la escala de Landis y Koch para K (10) que plantea:

- Excelente: $k \geq 0.80$ (asociación muy fuerte)
- Buena: $0.60 \leq k < 0.80$ (Fuerte)
- Moderada: $0.40 \leq k < 0.60$ (Mediana)
- Escasa: $0.20 < k < 0.40$ (Débil)
- Nula: $k < 0.20$ (no hay asociación)

Se empleó la prueba Ji-cuadrado de homogeneidad de K para determinar diferencias significativas entre las pruebas estudiadas.

Para la determinación de los indicadores básicos de las pruebas diagnósticas se siguió la metodología propuesta por Díaz *et al.* (9). Se consideró al EF como "prueba de Oro" y al C1, PIS y TCC como las "pruebas diagnósticas" o predictoras. Se construyeron las tablas 2x2 correspondientes y se consideraron casos positivos aquellos estudiantes que suspendían el C1, PIS o TCC y los que suspendían el EF como "enfermos". Lo opuesto se consideró para los casos negativos y los "sanos". De esta forma los indicadores quedaron representados como:

- 1.Sensibilidad (S): porcentaje del total de "enfermos" propuesto por el diagnosticador.
- 2.Especificidad (E): porcentaje del total de "sanos" propuesto por el diagnosticador.
- 3.Valor predictivo positivo (VPP): porcentaje del total de casos dados como "positivos" por el predictor, que suspendieron realmente el EF.
- 4.Valor predictivo negativo (VPN): porcentaje del total de casos dados como "negativos" que apro-

Tabla 1. Porcentaje de promoción esperada según la variable estudiada y resultados reales obtenidos en el examen final ordinario de MIR por cada año.

Curso escolar	C1	PIS	TCC	EF	N
2001/2002	57,8	44,2	67,9	68,8	554
2002/2003	49,2	44,2	63,8	68,2	437
2003/2004	57,8	48,9	62,4	72,9	505
2004/2005	56,5	42,5	65,7	69,7	492
2005/2006	69,2	46,9	62,6	62,7	535

Leyenda: C1: Corte evaluativo; PIS: Prueba Intrasesemestral; TCC: Trabajo de Control en Clases; EF: Examen Final Ordinario; N: Tamaño de muestra.

baron el EF.

Todos los cálculos y pruebas estadísticas se realizaron en los paquetes estadísticos SPSS y Epidat 3.0.

RESULTADOS

-Promoción estimada y real para el EF ordinario de MIR por cada curso académico

En la tabla 1 se muestran los porcentajes de aprobados estimados, según cada diagnosticador y los reales del EF de MIR. Según estos datos se determinó que el TCC fue en los cinco cursos analizados la prueba cuyos resultados de promoción se acercan más a los valores reales de aprobados en el EF de MIR. La PIS fue la que peores resultados mostró sobreestimando bastante el porcentaje real de desaprobados.

-Indicadores de asociación y del desempeño de las pruebas diagnósticas

El análisis de la Tabla 2 indica un coeficiente rs significativo en todos los casos ($p < 0.001$) mostrando así una asociación entre las notas del EF con cada uno de los resultados de los diagnosticadores (C1, PIS y TCC). Sin embargo, el TCC y

la PIS poseen los mayores valores de rs.

Se determinó K, un coeficiente que permite medir la fortaleza de esta asociación y que tiene en cuenta, además, la influencia del azar en los resultados obtenidos (10). Por otro lado, los indicadores del desempeño de las pruebas diagnósticas, también apuntaron nuevamente hacia la PIS y el TCC como mejores predictores de los resultados en el EF de MIR (Tablas 3 y 4).

DISCUSIÓN

Los resultados de promoción estimada y real obtenidos para el EF ordinario de MIR por curso académico, señalan al TCC como prueba parcial cuyos resultados de promoción se acercan más a los valores reales de aprobados en el EF de MIR. Algunos autores indican que la mejoría observada en el TCC es comprensible teniendo en cuenta que los estudiantes ganan experiencia frente a los exámenes en la medida que avanza el curso (11). Sin embargo, también se debe destacar que esta mejoría puede estar influida por la ayuda diferenciada que se presta por parte de profesores, monitores y alumnos ayudantes a los alumnos suspensos en la PIS.

Tabla 2. Valores del coeficiente de correlación de Spearman (rs) y Kappa de Cohen (K) según curso académico y variable analizada.

Curso escolar	rs *			K*		
	C1	TCC	PIS	C1	TCC	PIS
2001/2002	0,53	0,68	0,64	0,32	0,46	0,39
2002/2003	0,55	0,64	0,67	0,33	0,43	0,36
2003/2004	0,52	0,60	0,56	0,32	0,38	0,32
2004/2005	0,53	0,54	0,58	0,31	0,36	0,34
2005/2006	0,49	0,54	0,61	0,28	0,35	0,40

Leyenda: * Todos los valores son significativos con una probabilidad menor de 0,01
 C1: Corte evaluativo; PIS: Prueba Intrasestremstral; TCC: Trabajo de Control en Clases; EF: Examen Final Ordinario.

La asociación entre las notas del EF y cada uno de los diagnosticadores (C1, PIS y TCC), se observó con anterioridad en otros estudios realizados donde los resultados de los exámenes parciales correlacionaron mejor con los del EF de

Bioquímica (11, 12).

Según la escala de Landis y Koch, los valores de K encontrados en este estudio, a pesar de ser significativamente diferentes de cero ($p < 0,01$), reflejan una débil asociación entre los

Tabla 3. Indicadores del desempeño de las pruebas diagnósticas Sensibilidad y Especificidad, según curso académico y variable predictora analizada.

Curso escolar	Sensibilidad*			Especificidad*		
	C1	TCC	PIS	C1	TCC	PIS
2001/2002	75,1	68,8	90,2	72,7	84,5	59,8
2002/2003	80,6	72,7	92,8	63,1	84,7	61,4
2003/2004	75,2	77,4	86,9	70,1	77,2	62,1
2004/2005	72,5	64,6	88,6	69,1	79,6	56,0
2005/2006	60,0	63,0	84,5	77,6	79,9	65,7

Leyenda: * Los valores se presentan en por ciento, C1: Corte evaluativo; PIS: Prueba Intrasestremstral; TCC: Trabajo de Control en Clases; EF: Examen Final Ordinario.

Tabla 4. Indicadores del desempeño de las pruebas diagnósticas: valor predictivo positivo (VPP) y valor predictivo negativo (VPN) según curso académico y variable analizada.

Curso escolar	VPP*			VPN*		
	C1	TCC	PIS	C1	TCC	PIS
2001/2002	55,6	66,9	50,5	86,6	85,6	93,1
2002/2003	50,5	63,9	52,9	87,4	90,0	94,8
2003/2004	48,1	55,8	46,1	88,4	90,2	92,3
2004/2005	50,5	58,6	46,6	85,3	84,2	91,9
2005/2006	61,5	63,0	59,5	76,5	77,9	87,7

Leyenda: * Los valores se presentan en por ciento, C1: Corte evaluativo; PIS: Prueba Intrasesemestral; TCC: Trabajo de Control en Clases; EF: Examen Final Ordinario.

resultados esperados y los observados, esto coincide con lo encontrado previamente por Pedrosa *et al* (datos no publicados) para esta asignatura en la ELAM.

El hecho que la PIS detectara un mínimo de ocho de cada 10 estudiantes que desaprobaban el EF, se debe a que este examen parcial ha sobreestimado históricamente la cantidad de estudiantes suspensos en el EF. Por otro lado, los valores de VPN determinados para la PIS muestran que más del 90% de los estudiantes propuestos como posibles aprobados, realmente aprobaron el EF, lo que se explica si se tiene en cuenta que este examen parcial, además de ser más abarcador e integrador, evalúa conocimientos y habilidades que serán posteriormente empleados como base para los temas subsiguientes. Un estudiante que apruebe la PIS, debe haber vencido un nivel de conocimientos y habilidades tal que le resulten útiles en la comprensión de los temas que siguen.

En la literatura existen pocas investigaciones similares a la realizada en este trabajo. Un estudio realizado en la ELAM con la asignatura BCM mostró que la PIS como examen parcial, en comparación con la evaluaciones frecuentes, tenía también mejores indicadores como prue-

ba diagnóstica. De forma similar a lo observado en este trabajo, los VPP fueron bastante bajos en la mayoría de los cursos analizados (9). A pesar que las comparaciones entre ambas asignaturas son limitadas por las diferencias de contenidos y habilidades que el estudiante debe vencer, se puede decir que no existen evidencias que indiquen un comportamiento diferente entre BCM y MIR como el reportado por Lavín *et al.* (12).

Los resultados apuntan hacia la PIS y el TCC como mejores diagnosticadores del éxito o el fracaso en el EF de MIR. Sin embargo, considerando que los resultados del TCC se obtienen en las postrimerías del curso académico y que los del C1 y la PIS se tienen a mediados de semestre, es factible entonces decir que la mejor prueba diagnóstica es la PIS.

Aunque este examen sobreestima los desaprobados en el EF, sería mucho mejor dedicar atención diferenciada a un grupo superior de estudiantes dentro de los cuales estén los que con mayor probabilidad suspenderían el EF, que no hacerlo.

No obstante lo planteado antes, se debe puntualizar que la PIS es un examen que se realiza a mediados del curso. Esto dejaría muy poco

tiempo para el trabajo diferenciado con los estudiantes de alto riesgo académico. Otros autores han propuesto varios indicadores para medir riesgo académico de los estudiantes en las carreras médicas. Uno de los más sencillos es el cómputo de los resultados obtenidos en los preuniversitarios y las pruebas de ingreso (6, 13); sin embargo, esto es prácticamente inaplicable en la ELAM, por la heterogeneidad de sus educandos. Una alternativa factible sería realizar pruebas de nivel de entrada (PNE) al primer año de la carrera de medicina (6) como indicadores tempranos de este riesgo académico, pero en estos momentos no se cuentan con cuestionarios validados para los estudiantes de nuestro centro.

CONCLUSIONES

En todos los cursos existió una asociación significativa entre los resultados obtenidos en el examen

final de la asignatura Metabolismo Intermediario y su Regulación y los esperados, según el corte evaluativo, la prueba intrasemestral y el trabajo de control en clase. No obstante, los valores de K indican que tal asociación es leve, según la escala de Landis y Koch.

La prueba intrasemestral y el trabajo de control en clase fueron, además, las pruebas diagnósticas con mejores indicadores de Sensibilidad/ Valor Predictivo Negativo y de Especificidad/ Valor Predictivo Positivo, respectivamente. Detectar en menor tiempo más del 80% de los estudiantes que realmente suspenden el examen final se considera un elemento importante en la selección de la prueba intrasemestral como mejor predictor de los resultados finales de esta asignatura. ∞

BIBLIOGRAFIA

1. Díaz GR, Valdés de la RC, Hernández GS, Nápoles VA, Fajardo CB y Pedrosa AA. Valoración del nivel de conocimientos y habilidades con que ingresan los estudiantes a la carrera de medicina. *Rev Cubana Educ Med Super* 2001;15:172-9.
2. Valdés de la RC, Álvarez AN, Hernández GS, González GR, y Fajardo CB. Diseño del sistema de habilidades intelectuales en la asignatura Bioquímica I del primer año de la carrera medicina. *Rev Cubana Educ Med Super* 2001;15:117-26.
3. Vélez van MA, Roa GCN. Factores asociados al rendimiento académico en estudiantes de medicina. *Educ Med* [serie en Internet]. 2005;8:24-32. Disponible en: http://www.scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1575-18132005000200005&lng=en&nrm=iso, [Citado 11 de abril de 2007].
4. Boffill CM, González GOL, Cárdenas MN. Rendimiento docente de los estudiantes de medicina en la asignatura Bioquímica. *Medicentro* 1987;3:78-86.
5. Rosell PW. La enseñanza integrada de las Ciencias Médicas. *Rev Cubana Educ Med Super* 1998;12:45-8.
6. Bacallao GJ. Al rescate de las pruebas de nivel de entrada como predictores del rendimiento en la enseñanza médica superior *Rev Cubana Educ Med Sup* [serie en Internet] 1997;10. Disponible en: <http://bvs.sld.cu/revistas/ems/indice.html> [Citado 10 de Octubre de 2006].
7. Carrión PE. Validación de caracteres de ingreso como predictores del rendimiento académico en la carrera de medicina. *Rev Cubana Educ Med Sup* [serie en Internet] 2002;16. Disponible en: <http://bvs.sld.cu/revistas/ems/indice.html>. [Citado 15 de Octubre de 2006].
8. Hidalgo de PA, Gil LM, Rodríguez RE. Capacidad de atención y resultados docentes de estudiantes de primer año de medicina. *Rev Cubana Educ Med Super* 2001;15:273-8.
9. Díaz AD, Hernández FM, Serrano PK. Corte evaluativo y prueba intrasemestral como predictores de los resul-

tados finales en Biología Celular y Molecular. Panorama Cuba y Salud 2006;1:24-31.

10. Kraemer HC, Bloch DA. Kappa coefficients in epidemiology: an appraisal of a reappraisal. J Clin Epidemiol 1988;41:959-68.

11. Cruz GJ, Ribero MMR, Rubio BE y Lavín PM. Estudio de diferentes componentes del sistema de evaluación en Bioquímica en la Licenciatura en Enfermería Rev Hab Ciencias Méd [serie en Internet] 2002;1. Disponible en: <http://www.ucmh.sld.cu/rhab/> [Citado 9 de Octubre de 2006].

12. Lavín PM, Alé SO, Rivero MR y Corbelle QJ. Grados de asociación entre las evaluaciones frecuentes, parcial y final de Bioquímica en Medicina. Rev Hab Ciencias Méd [serie en Internet] 2002;1. Disponible en: <http://www.ucmh.sld.cu/rhab/> [Citado 9 de Octubre de 2006].

13. Rodríguez R, Díaz PA, Moreno M, Bacallao J. Capacidad predictiva de varios indicadores de selección para el ingreso a la carrera Medicina. Rev Cubana Educ Med Super 2000;14:128-135.

Frequent and partial exams as indicators of the final results of the subject Intermediary Metabolism and its Regulation

Abstract

Objectives: To determine the degree of association among the results of the frequent tests (evaluation cut or C1) and partial tests (inter-semester exam or ISE) and control work in class or EWC) with those from the Ordinary Final Exam (FE) of the subject Intermediary Metabolism and its Regulation (IMR) and evaluate through the calculation of basic indicators of the diagnostic test performance the ISE EWC and C1 value as those that diagnose the Ordinary exam results of this subject at LAMS.

Methods: A transversal comparative study that involved the C1 results was carried out on 2 523 second semester medicine students from September 2001 to July 2006.

Results: In all the courses a significant association among the results obtained in the FE and the expected ones was found according to C1, IMR, EWC, examinations. The negative predictive value and sensibility indicators were better than the ones in the EWC.

Conclusions: This work demonstrates that ISE in the Latin American Medical School is the best way to predict the final results in the subject Intermediary Metabolism and its Regulation.

Key words: Biochemistry, evaluation, learning, higher education.

Dirección para la correspondencia:

Lic. Dariel Díaz Arce. Escuela Latinoamericana de Medicina (ELAM), Departamento Bioquímica, Carretera Panamericana Km. 3 1/2, Sta. Fe, Playa, Ciudad de la Habana

E-mail: dariel@elacm.sld.cu