

Corte evaluativo y prueba intrasemestral como predictores de los resultados finales en Biología Celular y Molecular

Dariel Díaz Arce*

Martha Hernández Fernández**

Karel Serrano Patterson***

*Lic. en Bioquímica, Profesor Instructor, Departamento de Bioquímica, Escuela Latinoamericana de Medicina (ELAM).

**Lic. en Bioquímica, Master en Ciencias Químicas, Profesora Auxiliar, Departamento de Bioquímica, ELAM.

***Lic. en Química, Profesor Instructor, Departamento de Bioquímica, ELAM.
e-mail: dariel@elacm.sld.cu

Resumen

Objetivo: determinar la validez del Corte Evaluativo y la Prueba Intrasemestral como predictores de los resultados finales de la asignatura Biología Celular y Molecular, teniendo en cuenta los indicadores del desempeño de estas herramientas diagnósticas.

Métodos: Estudio comparativo transversal. Se emplearon los resultados del Corte, la Prueba Intrasemestral y del Examen Final tomados al azar de 24 grupos de estudiantes de primer año de la Escuela Latinoamericana de Medicina (ELAM), en los cinco cursos comprendidos entre los años 1999 y 2004. El grado de asociación se determinó por los valores de concordancia entre los resultados del Corte y la Prueba Intrasemestral con los resultados del Examen Final. Además, se determinó la sensibilidad, especificidad, valores predictivos positivo y negativo.

Resultados: La concordancia entre las variables analizadas predijo aproximadamente el 80% de los estudiantes que aprueban el Examen Final ordinario de Biología Celular y Molecular. La Prueba Intrasemestral logró detectar más resultados negativos que el Corte. Se propone una fórmula para predecir los resultados finales en la asignatura que tiene en cuenta los resultados históricos.

Conclusiones: Este trabajo muestra que la Prueba Intrasemestral en la ELAM es un buen indicador de los resultados en el Examen Final de Biología Celular y Molecular.

Palabras clave: Evaluación, aprendizaje, Bioquímica.

INTRODUCCIÓN

La disciplina Bioquímica ha sido tradicionalmente de difícil comprensión y asimilación por los estudiantes de las Ciencias Médicas. Entre los aspectos que incrementan la dificultad en el aprendizaje de esta materia están el nivel de abstracción que requiere el estudiante y su poca motiva-

ción al estudiarla (1,2).

La identificación de los alumnos con dificultades en la adquisición de estos conocimientos debe realizarse lo más temprano posible en aras de aplicar las medidas de atención personalizada pertinentes. De este modo, la atención a las diferencias individuales posibilita el cumplimiento de los objetivos instructivos y educa-

tivos de la asignatura (1, 2).

Estudios precedentes han analizado la utilidad de múltiples variables académicas, intelectuales y psicológicas, como predictoras de resultados docentes en las carreras médicas (3, 4, 5, 6, 7). Algunos de estos indicadores aplicados a la asignatura Biología Celular y Molecular (BCM), no fueron satisfactorios (3).

Estudios realizados por Lavín y cols. (8) y Cruz y cols. (9), aportaron evidencias que apoyan a los exámenes parciales como buenos predictores de los resultados finales en esta asignatura, aún en diferentes carreras.

Con el presente trabajo nos propusimos, determinar la validez del Corte Evaluativo (C1) y de la Prueba Intrasemestral (PIS) como predictores de los resultados finales de la asignatura BCM, a partir de los indicadores del desempeño de estas pruebas diagnósticas.

MATERIALES Y MÉTODOS

Se hizo un estudio comparativo transversal en el que se realizó un muestreo por conglomerados y se seleccionó al azar 24 grupos del primer año de cada uno de los cinco cursos académicos desarrollados en la ELAM en el periodo Septiembre 1999 - Julio 2004. El tamaño muestral fue de 3124 alumnos, representando alrededor del 35% del total de los estudiantes matriculados en estos cursos. Esto permitió realizar las estimaciones de intervalo de confianza de kappa (k) con una precisión absoluta de 0,035 y para la sensibilidad y especificidad menor del 3%, con una probabilidad del 95%. Los estudiantes que no tenían completas sus evaluaciones (101 alumnos) se consideraron resultados negativos.

Las variables PIS y C1 dispo-

nen de las categorías aprobado (A) y desaprobado (D) para la estimación de los valores de k , promoción y de los indicadores del desempeño de las pruebas diagnósticas. Para la estimación de los valores de calidad dichas variables toman los valores de Bien (B), Regular (R) y Mal (M), considerando las combinaciones de notas propuestas por la VAD No. 5 del año 1991 (10).

Todos los cálculos estadísticos se desarrollaron empleando los programas Microsoft Excel y EPIDAT 3.0.

Para el cálculo de la concordancia y de los indicadores estadísticos básicos se emplearon las tablas 2 x 2 correspondientes y se consideró el Examen Final (EF) ordinario de BCM como "Prueba de Oro".

Se consideraron "enfermos" aquellos estudiantes que desaprobaron el EF y como "sanos" los que lo aprobaron. En este sentido la sensibilidad y la especificidad quedaron definidas como:

-Sensibilidad: Por ciento de los estudiantes que desaprobaron el EF y fueron propuestos como desaprobados en el C1 o PIS.

-Especificidad: Por ciento de los estudiantes que aprobaron el EF y fueron propuestos como aprobados en el C1 o PIS.

Asimismo, se definieron los valores predictivos negativo y positivo (VPN y VPP, respectivamente) como la probabilidad de que un estudiante, dado como posible aprobado o posible sus-

penso según el diagnosticador, llegue a alcanzar realmente tal categoría en el EF.

Los valores de sensibilidad, especificidad, VPN y VPP, se expresaron en por ciento.

La determinación de la concordancia se realizó mediante la cuantificación de k . Para determinar la significación de este valor se empleó la escala de Landis y Koch (11) en la que:

-Excelente: $k \geq 0,80$ (asociación muy fuerte)

-Buena: $0,60 \leq k < 0,80$ (fuerte)

-Moderada: $0,40 \leq k < 0,60$

(moderada)

-Escasa: $0,20 < k < 0,40$ (débil)

-Nula: $k < 0,20$ (no hay asociación).

Se empleó la prueba de Chi cuadrado de homogeneidad de k para determinar las diferencias entre los valores obtenidos para la PIS y para el C1.

Teniendo en cuenta la probabilidad histórica de aprobar o suspender según el C1 o la PIS, se diseñó una estrategia para obtener una fórmula que permitiera estimar los resultados de promoción y calidad del EF ordinario en cursos subsecuentes.

RESULTADOS

1. Determinación de la concordancia y comparación de resultados.

La Tabla 1 muestra los valores de k obtenidos para cada predictor. Ambos valores, k_{PIS} y k_{C1} , fueron significativos y dife-

Tabla 1. Concordancia entre los resultados del C1 y la PIS con los resultados en el Examen Final Ordinario de BCM.

Prueba		Examen Final		k [IC, 95,0 %]
		D	A	
C1	D	601	388	0,42 [0,38-0,45]
	A	406	1729	
PIS	D	738	368	0,55 [0,51-0,58]
	A	269	1749	

D: estudiantes desaprobados, **A:** estudiantes aprobados
k: valor de Kappa (Intervalo de confianza 95%)

rentes entre sí ($p < 0,001$), mostrando la PIS una mejor concordancia con los resultados del EF, aunque ambos fueron considerados moderados según la escala de Landis y Koch (11).

2. Determinación de los indicadores básicos del desempeño de las pruebas diagnósticas.

En las Tablas 2 y 3 se muestran los resultados de la sensibilidad, especificidad y valores predictivos, empleando como diagnósticos al C1 y a la PIS, para cada curso académico.

Según aquí se refleja, se encontró para la PIS en casi todos los aspectos, mejores indicadores de pruebas diagnósticas que para el C1. Se comportaron de modo similar en la especificidad. La PIS detectó mayor número de estudiantes que suspendieron el EF, siendo la sensibilidad superior al 70 % en cuatro de los cinco cursos analizados, mientras que para el

C1, este valor solo se rebasó en un curso académico.

3. Estimación de los resultados finales a partir de los resultados históricos.

3.1. Estimación de la Promoción.

En la Tabla 4, se reflejan los

estudiantes propuestos como posibles aprobados (a), lo que representa la fracción $a/(a+b)$ del total de los propuestos. Esto a su vez se correspondió con el VPN/100 a quien se denominó VPN'. Por otro lado, se tiene que de los que se propusieron como suspensos por la PIS o el C1, lograron aprobar el EF la cantidad c cuya fracción del total propuestos sería $c/(c+d)$. Para esta última fracción, si se multiplica el numerador y el denominador por el factor d, se puede entonces llegar a la relación $cx/(c+d)xd$, donde $d/(c+d)$ se corresponde con el VPP/100 a quien se llamará VPP'.

De esta forma se obtiene la cantidad aproximada de posibles aprobados en el EF a partir de la siguiente fórmula:

Tabla 2. Corte Evaluativo como predictor de resultados en el Examen Final ordinario de BCM.

Curso	Corte Evaluativo			
	E	S	VPN	VPP
1999/2000	84,7	48,6	84,5	48,9
2000/2001	73,1	74,1	81,5	63,7
2001/2002	87,8	60,0	83,5	68,0
2002/2003	78,9	56,0	71,6	65,4
2003/2004	89,7	39,4	85,0	50,0
Todos	81,7	59,7	81,0	60,8
IC (95 %)	81,6 - 81,7	59,6 - 59,8	81,0 - 81,0	60,7 - 60,8

VPN: Valor Predictivo Negativo; **VPP:** Valor Predictivo Positivo;
E: Especificidad, **S:** Sensibilidad, **IC (95 %):** Intervalo de confianza con una probabilidad del 95 %.

Posibles aprobados

$$=VPN \cdot x(B+R) + VPP \cdot x(c/d) \cdot x(M)$$

(Fórmula 1)

Donde B+R representa el total de los posibles aprobados según el corte o la PIS y M el total de suspensos en el curso que se quiera analizar; c/d se determina por los valores históricos, según la tabla anterior.

La aplicación de esta fórmula a los resultados del primer semestre del curso 2004 -2005 arrojó los resultados representados en la Figura 1.

3.2. Estimación de la calidad

Para ello se procedió de forma similar a la anterior, como se muestra en la Tabla 5.

La calidad fue determinada por el número de estudiantes que obtienen evaluación de Bien (B) (notas de 4 ó 5 puntos) en el EF, entre el número total de estudiantes que aprobaron dicho examen. En este caso está dado por la fórmula:

$$\text{Calidad} = (a+d+g) / (\text{Total de aprobados}) \times 100$$

(Fórmula 2)

Los valores a, d y g de la

Tabla 3. Prueba Intrasemestral como predictor de resultados en el Examen Final Ordinario de BCM.

Prueba Intrasemestral				
Curso	E	S	VPN	VPP
1999/2000	81,4	78,2	92,3	56,6
2000/2001	79,0	77,7	85,6	68,7
2001/2002	84,9	76,8	88,1	71,4
2002/2003	94,5	50,7	73,6	86,4
2003/2004	81,7	79,7	92,0	60,5
Todos	82,6	73,3	86,7	66,7
IC (95 %)	82,6 - 82,7	73,2 - 73,4	86,6 - 86,7	66,7 - 66,8

VPN: Valor Predictivo Negativo; **VPP:** Valor Predictivo Positivo; **E:** Especificidad, **S:** Sensibilidad, **IC (95 %):** Intervalo de confianza con una probabilidad del 95 %.

Fórmula 2 se pueden calcular a partir de los resultados históricos del siguiente modo:

$$\alpha = a / (\text{Número total de estudiantes con B en C1 o PIS})$$

$$\beta = d / (\text{Número total de estudiantes con R en C1 o PIS})$$

$$\gamma = g / (\text{Número total de estudiantes con M en C1 o PIS})$$

Con estos resultados históricos, se puede entonces calcular el número de estudiantes que se esperan obtengan

calificación B en el EF, a partir de los resultados del C1 o la PIS, según la Fórmula 3.

Estudiantes con $B' = (a)(B) + (b)(R) + (g)(M)$ (Fórmula 3)

Donde B, R y M representan la cantidad de estudiantes con esos resultados en el C1 o la PIS del año que se analice.

Luego la calidad esperada quedaría entonces:

$$CC = (B' / \text{Total de estudiantes aprobados según Fórmula 1}) \times 100. \quad (\text{Fórmula 4})$$

Empleando esta fórmula se obtienen los resultados mostrados en la Figura 2.

DISCUSIÓN

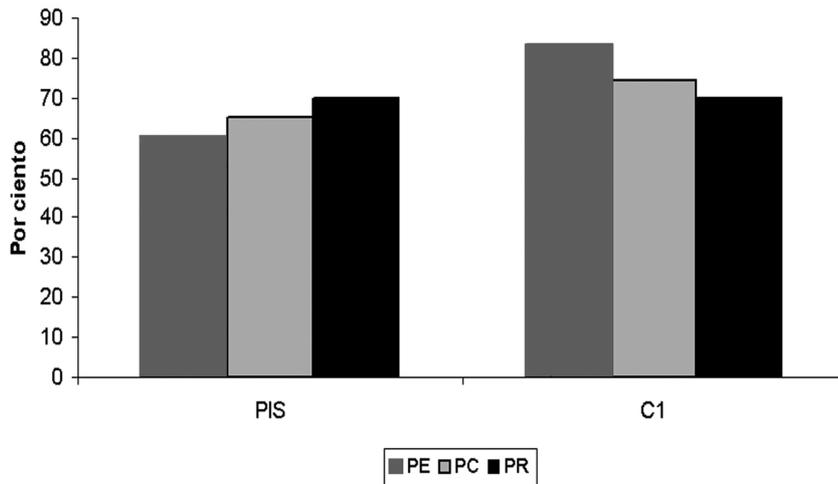
Estudios precedentes han puesto de manifiesto lo difícil de obtener predictores confia-

Tabla 4. Estimación de la promoción final en ordinario.

Variable (C1 ó PIS)	Examen Final		
	A	D	TOTAL
A	a	b	a+b
D	c	d	c+d
TOTAL	a+c	b+d	a+b+c+d

D: estudiantes desaprobados, **A:** estudiantes aprobados

Figura 1. Estimación de los resultados de promoción en el Examen Final Ordinario de BCM.



PE: Promoción Estimada de forma tradicional; **PC:** Promoción Calculada empleando la Fórmula 1; **PR:** Promoción Real obtenida en el Examen Final.

bles en la estimación de los resultados en el EF de BCM (3). Normalmente, los resultados de las evaluaciones del C1 son empleados en las instituciones de Educación Médica Superior para dar un estimado "bruto" de los resultados en los exámenes finales de cada asignatura. En este trabajo, se presentan evidencias que apuntan hacia la PIS como un buen indicador para tales fines.

En este estudio se decidió emplear el índice k como medida de asociación entre las variables analizadas, ya que este toma en cuenta la posible influencia del azar en los resultados observados (11). Aunque en ambos predictores el valor de k indica una asociación moderada, según la escala de Landis y Koch (11) se debe resaltar que para la PIS esta relación fue significa-

tivamente superior ($p < 0,001$). Lo anterior estuvo apoyado por la mayor parte de los indicadores del desempeño de las pruebas diagnósticas presentados para cada curso en la Tabla 2. No obstante, lo más interesante estuvo en que la sensibilidad de la PIS permitió detectar aproximadamente a tres de cada cuatro estudiantes que suspenden el EF de la asignatura. Algunos autores consideran que para

las pruebas diagnósticas de este tipo, una sensibilidad superior al 60 % puede considerarse de buena (12). En este caso se considera el valor de 73 % de muy bueno ya que detectando a tiempo a los estudiantes con alto riesgo de suspender, se pueden diseñar estrategias de trabajo encaminadas a elevar su preparación antes del EF.

Una investigación realizada en la ELAM en el año 2003 con la asignatura Metabolismo Intermediario y su Regulación, arrojó resultados de concordancia también moderados para ambos predictores (datos no publicados).

Aunque en este caso no hubo diferencias significativas entre los valores de k_{PIS} y k_{C1} , el cálculo posterior de los valores de los indicadores de las pruebas diagnósticas apoyó también a la PIS como mejor predictor de resultados en el EF de esta otra asignatura. Salvando las posibles diferencias que puedan surgir del

Tabla 5. Estimación de la calidad final en el Examen Final ordinario.

Variable (PIS ó C1)	Examen Final			Subtotal
	B	R	M	
B	a	b	c	a+b+c
R	d	e	f	d+e+f
M	g	h	i	g+h+i
Subtotal	a+d+g	b+e+h	c+f+i	Total

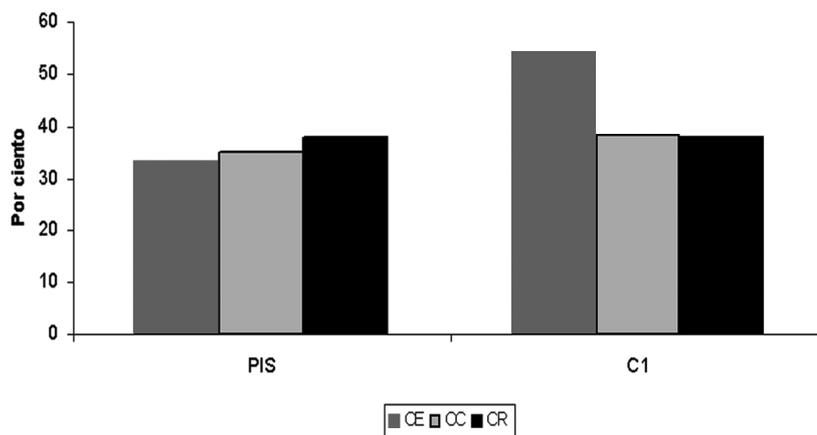
B, R, M: Notas de Bien, Regular y Mal respectivamente.

grado de complejidad de ambas materias, lo anterior sugiere un comportamiento similar para ellas, lo que al parecer no ocurre del mismo modo en otras instituciones de Educación Médica Superior (8).

En cuanto a los VPN (Tablas 2 y 3) la PIS también presentó los mejores resultados con indicadores bastante altos en cuatro de los cinco cursos analizados. Aunque los valores predictivos positivos en la PIS no son muy elevados, sí cabe destacar que en todos los cursos han sido superiores a los del C1, indicando que aproximadamente dos de cada tres estudiantes propuestos como suspensos por esta prueba salieron mal en el EF.

Otras investigaciones recientes han aportado evidencias que también sustentan los resultados de este trabajo. Lavín y cols. (8) plantean que en la asignatura BCM de la carrera Medicina, existe una buena correlación entre los resultados de la PIS y los del EF. Esta correlación es superior a la observada con los resultados de las actividades frecuentes de seminarios y prácticas. Un estudio similar en la carrera Enfermería, mostró una mejor correlación, entre los resultados de los trabajos de control en clase y los del EF, que la que se obtiene con los informes de esta asignatura (9).

Figura 2. Estimación de la calidad en el Examen Final Ordinario de BCM.



CE: Calidad Estimada de forma tradicional; **CC:** Calidad Calculada según la fórmula 4; **CR:** Calidad Real obtenida en el EF.

Los resultados expuestos aquí son, en cierta medida lógicos, si se considera que el nivel de información que un estudiante debe dominar para aprobar en un examen parcial es muy superior al de una evaluación frecuente.

Se debe señalar que este trabajo empleó un punto de corte arbitrario para el cálculo de los indicadores estadísticos de las pruebas diagnósticas. Algunos autores han planteado que estas decisiones pueden generar una imagen distorsionada de la relevancia relativa del predictor (13). Sin embargo, aquí se emplearon los puntos de corte que definen los aprobados y desaprobados, según las reglamentaciones de la VAD No.5 de 1991 (10).

Las Figuras 1 y 2 indican que, al menos para el curso analizado 2004 - 2005, las

fórmulas propuestas en este trabajo para la estimación de los resultados finales de promoción y calidad en ordinario de la asignatura BCM, dieron resultados bastante cercanos al valor real. Procedimientos similares no se han encontrado en la literatura.

No obstante, aunque los resultados aquí expuestos son alentadores, es pertinente aclarar que es necesario profundizar en ellos para garantizar su validación, por lo que solo se deben considerar como preliminares. Además, se estima que los resultados de las fórmulas 1 y 4 pueden estar sesgados por múltiples factores de difícil medición, ya sean propios del profesor, propios del estudiante, derivados de la interacción de ambos o por factores externos al proceso de enseñanza aprendizaje. En este sentido, algunos estudios

han indicado que la capacidad de atención de los alumnos, la calidad de los textos para el estudio individual, así como su preparación previa antes de entrar en la carrera medicina, pueden ser factores que afectan la promoción general en esta disciplina (3, 5, 6, 7, 14).

Algunas de estas variables pudieran ser controladas en la ELAM por los resultados del curso Pre-médico y en parte por la caracterización psicopedagógica de los alumnos, lo que contribuiría a una mejor y más abarcadora estimación de los resultados en la asignatura estudiada.

Por otra parte, si se detecta

tempranamente en un alumno, dificultades con el aprendizaje de la asignatura y se desarrolla un trabajo dirigido a superar sus deficiencias, esto atendería de hecho contra cualquier fórmula matemática que intente predecir los resultados finales.

Un análisis minucioso de lo anterior pudiera aportar un factor de corrección para dichas fórmulas que incluya algunas de las variables que afecten positiva o negativamente al proceso docente.

CONCLUSIONES

El análisis de los resultados de la PIS y C1 de BCM duran-

te los 5 cursos anteriores mostró que ambos son buenos predictores de los posibles aprobados en el EF, lo que se evidencia en valores de especificidad y VPN por encima del 80 %. Sin embargo, la concordancia y la sensibilidad han sido mejores en el caso de la PIS, con valores, esta última, por encima del 70 % en cuatro de los cinco cursos analizados. En ambos casos, los VPP fueron similares. Estos resultados pudieran emplearse para predecir los resultados finales en el EF ordinario de BCM, aunque se requiere profundizar en este estudio.

BIBLIOGRAFÍA

1. Rosell Puig W. *La enseñanza integrada en las Ciencias Médicas*. Rev Cubana Educ Med Sup 1998;12(2):45-48
2. Díaz G R, Valdés de la Rosa C, Hernández G S, Nápoles V A, Fajardo C B y Pedrosa A A. *Valoración del nivel de conocimiento y habilidades con que ingresan los estudiantes a la carrera medicina*. Rev Cubana Edu Med Sup 2001;15(2):172-179
3. Bacallao G J. *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 [Citado 10 de Octubre de 2005];10(1): [Aprox. 12 p]. Disponible en: <http://bvs.sld.cu/revistas/ems/indice.html>
4. Bacallao G J, Parapar JM, Roque MY, Bacallao G J. *La modelación jerárquica y efectos de grupo en la predicción del rendimiento académico*. Educ Med Sup [serie en Internet] 2004 [Citado 15 de Octubre de 2005];18(2):[Aprox.10 p] Disponible en: <http://bvs.sld.cu/revistas/ems/indice.html>
5. Carrión P E. *Validación de caracteres de ingreso como predictores del rendimiento académico en la carrera de medicina*. Rev Cubana Educ Med Sup [serie en Internet] 2002 [Citado 15 de Octubre de 2005]; 16(2):[Aprox.10 p] Disponible en: <http://bvs.sld.cu/revistas/ems/indice.html>.
6. Roque F R, Bacallao G J, Díaz R P, Morejón L M. *Valor predictivo de algunos criterios de selección para el ingreso a la carrera de medicina*. Rev Cubana Educ Med Sup 2000; 14 (1): 17-25.
7. Hidalgo de Paz A, Gil L M, Rodríguez R E. *Capacidad de atención y resultados docentes de estudiantes de primer año de Medicina*. Rev Cubana Educ Med Super. 2001;15(3):273-276.
8. Lavín P M, Alé S O, Rivero MM R y Corbelle Q J. *Grados de asociación entre las evaluaciones frecuentes, parcial y final de Bioquímica en Medicina*. Rev Habanera Ciencias Médicas [serie en Internet] 2002 [Citado el 9 de Octubre de 2005]; 1(3):[Aprox.10 p] Disponible en: <http://www.ucmh.sld.cu/rhab/>
9. Cruz G J, Ribero MM R, Rubio B E y Lavín P M. *Estudio de diferentes componentes del sistema de evaluación en Bioquímica en Licenciatura en Enfermería*. Rev Habanera Ciencias Médicas [serie en Internet] 2002 [Citado el 9 de Octubre de 2005];

1(3):[Aprox.10 p] Disponible en: <http://www.ucmh.sld.cu/rhab/>

10. MINSAP. Instrucción v. a. d. No. 5. Ciudad de la Habana, 28 de octubre de 1991.

11. Kraemer HC, Bloch DA. Kappa coefficients in epidemiology: an appraisal of a reappraisal. *J Clin Epidemiol* 1988; 41: 959 - 968

12. Pérez G JC. Capacidad de trabajo mental. Una nueva propuesta metodológica para su estudio. *Ciencias Pedagógicas*. [serie en Internet] 2000 [Citado el 11 de noviembre de 2006]; 1(3):[Aprox.10 p] Disponible en: <http://cied.rimed.cu/revista/13/articulos/a1r3capacidad.html>

13. Bacallao G J. Las curvas ROC y las medidas de detectabilidad para la validación de predictores del rendimiento docente. *Rev Cubana Educ Med Sup* [serie en Internet] 1997 [Citado 10 de Octubre de 2005];10(1): [Aprox. 12 p]. Disponible en: <http://bvs.sld.cu/revistas/ems/indice.html>

14. Lavín P M, Alé S O, Rubio B E, Rivero MM R, Corbelle Q J. Influencia de tres componentes del proceso docente sobre el examen de Biología Celular y Molecular. *Rev Habanera Ciencias Médicas* [serie en Internet] 2003 [Citado el 11 de noviembre de 2006]; 2(6):[Aprox.10 p] Disponible en: <http://www.ucmh.sld.cu/rhab/>

Abstract

Objective: To determine the validity of the Academic performance report and the Mid-term test as predictors of the final results in the Cellular and Molecular Biology subject, considering the performance indicators of these diagnostic tools.

Methods: Transversal comparative study. The results of all, the Academic performance report, the Mid-term test and the Final exam were used. They were taken at random from 24 groups of first year students in the Latin American Medical School (ELAM in Spanish) during five courses from 1999 to 2004. The association degree was determined by the values of correspondence between the results of both the Academic Performance Report and the Mid-term test, and those of the Final exam. Sensitivity, specificity, and positive or negative predictive values were also determined.

Results: The correspondence among the variables that were analyzed predicted about the 80% of students that pass the ordinary final exam of Cellular and Molecular Biology. The Mid-term test detected more negative results than the Academic performance report. A formula to predict the final results in the subject is proposed taking into account its historical results.

Conclusions: This study shows that the Mid-term test administered in ELAM is a good predictor of the results in the Final exam of the subject Cellular and Molecular Biology.

Key words: Evaluation, learning, Biochemistry.