

Discriminación y calibración de cuatro escalas de calificación del síndrome de disfunción orgánica múltiple. Estudio multicéntrico

Dr. Luis David Sánchez Velázquez,* Dra. María Elena Reyes Sánchez,[†] Dra. Dulce María D'Ector Lira,[‡] Dr. Luis Carlos Bonilla Rivera,[§] Enf. Angélica González González,^{||} Enf. María de Lourdes Magdaleno Padilla,[¶] Enf. María Guadalupe González Vega**

RESUMEN

Objetivo: Analizar cuatro escalas para la evaluación del síndrome de disfunción orgánica múltiple en pacientes críticos mexicanos.

Diseño: Estudio multicéntrico, prolectivo.

Lugar: Cuatro UCIs de hospitales de tercer nivel de la Ciudad de México.

Pacientes: Se incluyó a 391 pacientes (152 mujeres, 239 hombres, edad media 56.9 + 18.9 años).

Intervenciones: Ninguna.

Mediciones y resultados principales: Se determinó las escalas de LODS, MODS, SOFA y Bruselas y se calculó las áreas bajo la curva y el modelo de calibración. La estancia fue de 3.95 + 4.3 días (rango 1-33 días). Las áreas bajo la curva fueron: SOFA 90.0 + 2.4, Bruselas 89.3 + 2.4, LODS 89.1 + 2.7 y MODS 87.2 + 3.2. Los modelos de calibración fueron: Bruselas, $p = 0.182$; MODS, $p = 0.005$; LODS, $p = 0.003$; y SOFA, $p = 0.002$.

Conclusión: Aunque la diferencia entre los sistemas de evaluación no fue significativa proponemos el uso de la escala de Bruselas.

Palabras clave: Falla orgánica múltiple, medidas pronósticas, índice de severidad de la enfermedad, sepsis.

SUMMARY

Objective: To analyze four scores for the evaluation of the multiple organ dysfunction syndrome in Mexican critically ill patients.

Design: Multicenter, projective study.

Setting: Four ICUs of tertiary hospitals, Mexico City.

Patients: A total of 391 patients were enrolled (152 women, 239 men, mean age 56.9 + 18.9 years).

Interventions: None.

Measurements and main results: LODS, MODS, SOFA and Bruselas score were determined and the areas under the curve and model c calibration were performed. ICU stay was 3.95 + 4.3 days (range 1-33 days). The areas under the curve were: SOFA 90.0 + 2.4, Bruselas 89.3 + 2.4, LODS 89.1 + 2.7 and MODS 87.2 + 3.2. The models' calibration were: Bruselas ($p = 0.182$), MODS ($p = 0.005$), LODS ($p = 0.003$) and SOFA ($p = 0.002$).

Conclusion: Although the difference between scoring systems was not significant we propose the use of Bruselas score.

Key words: Multiple organ failure, outcome measures, severity or illness index, sepsis.

* Médico adscrito matutino a la Unidad de Terapia Intensiva del Hospital Ángeles de las Lomas.

[†] Médico adscrito vespertino a la Unidad de Terapia Intensiva Posquirúrgica del Hospital General del Centro Médico Nacional "La Raza", IMSS.

[‡] Jefa de la Unidad de Terapia Intensiva, Hospital Ángeles de las Lomas.

[§] Jefe de la Unidad de Terapia Intensiva Posquirúrgica Cardiovascular, Centro Médico Nacional La Raza, IMSS.

^{||} Jefe de Enfermería Vespertina de la Unidad de Terapia Intensiva, Instituto Nacional de Cancerología, SSA.

[¶] Enfermera Especialista de la Unidad de Terapia Intensiva, Instituto Nacional de Cancerología, SSA.

** Enfermera de la Unidad de Terapia Intensiva, Instituto Nacional de Cancerología, SSA.

A inicios de los 70 surgió el concepto de falla orgánica múltiple como la principal causa de morbilidad y mortalidad en los enfermos admitidos a la unidad de terapia intensiva (UTI).¹ Se estima que en Norteamérica el 15% de los enfermos hospitalizados en la UTI desarrollan el síndrome de disfunción orgánica múltiple,² el cual es responsable del 80% de todas las muertes en la UTI.³

En 1992 se publicó un consenso⁴ en el que se definió el síndrome de disfunción orgánica múltiple como la disfunción fisiológica progresiva de dos o más sistemas después de una amenaza aguda a la homeostasis sistémica. Los factores de riesgo son diversos e incluyen infección, condiciones inflamatorias no infecciosas, isquemia, exposición a toxinas y activación inmunológica.

Han surgido numerosos estudios donde se estudia la disfunción orgánica múltiple, sin embargo, cada uno de ellos propone un sistema de calificación que contiene rubros difíciles de cuantificar, o consideran la disfunción como un evento dicotómico presente/ausente como la escala propuesta por Knaus.⁵ Sin embargo, en las últimas publicaciones se encuentra una tendencia a proponer escalas con mayor fundamento estadístico y de mayor pragmatismo clínico.

En 1995, Bernard y cols.⁶ propusieron una escala para cuantificar la disfunción orgánica múltiple, Bruselas score, para estudios clínicos y la práctica clínica (cuadro I). El sistema se aplicó en forma re-

trospectiva en una base de datos, sin embargo, no reportaron su calibración y discriminación.

En el mismo año apareció el Múltiple Organ Dysfunction Score (MODS),⁷ escala para medir la gravedad del síndrome de disfunción orgánica múltiple (SDOM) diseñada a partir de una revisión de la literatura y la evaluación en una unidad de terapia intensiva (UTI) polivalente de un hospital de tercer nivel. En el grupo de desarrollo de la escala (n= 336 enfermos) el área bajo la curva de receptor operativo fue de 93.6% y en el grupo de validación (n= 356 enfermos) de 92.8% (cuadro II). Desafortunadamente, los autores no realizaron la prueba de bondad de ajuste para conocer la calibración del sistema.

En 1996, Vincent et al.⁸ publicaron la escala Sepsis-related Organ Failure Assessment (SOFA) realizada en un consenso realizado en París en octubre de 1994 por la Sociedad Europea de Medicina de Cuidado Intensivo (ESICM) (cuadro III). Dicha escala tuvo buena correlación con la tasa de mortalidad en 1643 enfermos sépticos en una muestra retrospectiva, sin embargo, no publicaron la discriminación y calibración del sistema.

Le Gall et al.⁹ en el mismo año, publicaron un sistema más sofisticado, Logistic Organ Dysfunction System (LODS) desarrollado sobre las bases de regresión logística múltiple. En su grupo de desarrollo (n= 10,547) su discriminación fue de 84.3% y en el grupo de validación (n= 2,605) fue del

Cuadro I. Escala de calificación del síndrome de disfunción orgánica múltiple Bruselas.

Órgano	Normal	Leve	Moderada	Disfunción orgánica clínicamente significativa	
				Grave	Extrema
Cardiovascular (presión sistólica)	> 90	< 90	< 90	< 90	< 90
Pulmonar (PaO ₂ /FiO ₂)	> 400	respuesta a líquidos 301-400	sin respuesta a líquidos 201-300 Lesión pulmonar aguda	pH < 7.3 101-200 SIRA	pH < 7.2 < 100 SIRA grave
SNC (Glasgow)	15	13-14	10-12	6-9	< 5
Coagulación (plaquetas)	> 120	81-120	51-80	21-50	< 20
Renal (creatinina mg/dL)	< 1.5	1.5-1.9	2.0-3.4	3.5-4.9	> 5
Hepático (bilirrubina mg/dL)	< 1.2	1.2-1.9	2.0-5.9	6.0-11.9	> 12

Cuadro II. Escala de calificación del síndrome de disfunción orgánica múltiple (MODS).

Órgano	0	1	2	3	4
Cardiovascular (PAR)*	< 10	10.1-15.0	15.1-20.0	20.1-30.0	> 30.0
Pulmonar (PaO ₂ /FiO ₂)	> 300	226-300	151-225	76-150	< 75
SNC (Glasgow)	15	13-14	10-12	7-9	< 6
Coagulación (plaquetas)	> 120	81-120	51-80	21-50	< 20
Renal (creatinina mg/dL)	< 1.13	1.14-2.26	2.27-3.96	3.97-5.65	> 5.66
Hepático (bilirrubina mg/dL)	< 1.17	1.18-3.5	3.6-7.0	7.1-14.0	> 14.1

Cuadro III. Escala de calificación del síndrome de disfunción orgánica múltiple relacionado a sepsis (SOFA).

Órgano	1	2	3	4
Cardiovascular (hipotensión)	PAM < 70 torr	Dopamina < 5 ó dobutamina (cualquier dosis)	Dopamina > 5, epinefrina < 0.1 ó norepinefrina < 0.1	< 90 pH < 7.3
Pulmonar (PaO ₂ /FiO ₂)	< 400	< 300	< 200 Con ventilador	< 100 Con ventilador
SNC (Glasgow)	13-14	10-12	6-9	< 6
Coagulación (plaquetas)	< 150	< 100	< 50	< 20
Renal (creatinina mg/dL) ó uremis	1.2-1.9	2.0-3.4	3.5-4.9 o < 500 mL/día	> 5.0 o < 200 mL/día
Hepático (bilirrubina mg/dL)	1.2-1.9	2.0-5.9	6.0-11.9	> 12.0

85.0%, y su calibración fue muy buena, $p=0.21$ y $p=0.50$, respectivamente.

El propósito del estudio es determinar la discriminación y calibración de cada una de las cuatro escalas ya referidas y encontrar la más sensible y específica en población mexicana.

PACIENTES Y MÉTODOS

Se trata de un estudio multicéntrico con propósito descriptivo, de procedimiento, con asignación obser-

vacional, con dirección longitudinal, con colección prolectiva de los datos y con componentes homodémicos, realizado en cuatro unidades de terapia intensiva: una polivalente de un hospital general privado, una terapia traumatológica, una oncológica y otra de manejo posoperatorio cardiovascular. Se excluyeron del estudio aquellos enfermos hospitalizados en la unidad por menos de 24 horas y aquellos en los que no se consiguió la información completa.

Se realizó capacitación del personal encargado de la captura de la información para obtener una

adecuada reproducibilidad del método en la validación clínica del estudio. En este trabajo no se investigó la concordancia inter-observador.

Se consignaron variables demográficas, calificaciones APACHE II, LODS, MODS, SOFA y Bruselas a las 24 horas del ingreso a la unidad y estado vital al egreso hospitalario.

El análisis consistió de estadística descriptiva para las variables demográficas y de estadística inferencial para el cálculo del área bajo la curva de receptor operativo y su error estándar por el método de Dorfman y Alf (discriminación) y la bondad de ajuste se realizó con la prueba de Hosmer-Lemeshow (calibración). Se consideró significativa una $p < 0.05$. Además, se realizó comparación entre áreas bajo la curva para determinar la más discriminativa usando el paquete estadístico SPSS v 10.

RESULTADOS

Se reclutaron 391 enfermos: 183 enfermos del Hospital Ángeles de las Lomas, 120 del Centro Médico Nacional La Raza, IMSS, 49 del Instituto Nacional de Cancerología y 39 del Hospital de Urgencias Traumatológicas, IMSS (cuadro V). La edad promedio fue de $56.9 + 18.9$ (15-96 años) con una mediana de 59 años. Hubo 152 mujeres (38.9%). La es-

tancia hospitalaria promedio fue de $13.5 + 9.8$ días (1-59 días) con una mediana de 11 días, y la estancia promedio en la UTI fue de $3.8 + 2.7$ días (1-33 días) con una mediana de 3 días. La mortalidad del grupo total fue de 9.1%. La calificación APACHE II promedio fue de $15.1 + 7.9$ puntos (3-26) con una mediana de 18.

Los puntajes de las escalas fueron muy similares (cuadro VI). Asimismo, las áreas bajo la curva ROC fueron parecidas, con sus intervalos de confianza al 95% empalmados (cuadro VII y figura 1), sin tener diferencia estocástica entre ellas. Sin embargo, en la calibración la única escala con significancia estadística fue Bruselas.

DISCUSIÓN

Las ventajas del empleo de las escalas de evaluación de la disfunción orgánica múltiple son:

1. Mejorar el entendimiento de la historia natural de la disfunción y la interrelación entre la falla de órganos.
2. Evaluar el efecto de nuevas terapias en el curso de la disfunción orgánica.
3. Vigilar la tendencia al deterioro o mejoría de los enfermos en forma individual.

Cuadro IV. Escala de calificación logística del síndrome de disfunción orgánica múltiple (LODS).

Órgano	5	3	1	0	1	3	5
Cardiovascular							
FC	< 30 ó	-	-	30-139 y	> 140 ó	-	-
PAS	< 40	40-69	70-89	90-239	240-269	> 270	-
Pulmonar en ventilación mecánica (PaO ₂ /FIO ₂)		< 150	>150	Sin ventilación mecánica			
SNC	3-5	6-8	9-13	14-15			
Glasgow							
Coagulación							
Leucocitos	-	< 1.0	1.0-2.4 ó	2.5-49.9 y	> 50	-	-
Plaquetas	-	-	< 50	> 50	-	-	-
Renal							
Urea	-	-	-	< 36	36-59	60-119	> 120
BUN	-	-	-	< 17	17-27.9	28-55.9	> 56
Creatinina o uresis	< 0.5	0.5-0.74	-	< 1.2 0.75-9.99	1.2-1.59	> 1.6 > 10	-
Hepático							
Bilirrubina	-	-	-	< 2.0 y	> 2.0 ó	-	-
TP	-	-	< 25%	< 3 (> 25%)	> 3	-	-

Cuadro V. Demografía de las poblaciones estudiadas. Las desviaciones estándar entre paréntesis.

Variable	HAL	CMNLR	INCan	HUT	Total
n	183	120	49	39	391
Edad	59.3 (18.2)	59.6 (8.6)	54.1 (16.6)	49.4 (22.4)	56.9 (18.9)
Género femenino (%)	28.2	64.1	48.9	23	41.6
APACHE II	6.2 (4.4)	4.5 (2.1)	15.1 (7.9)	14.9 (4.7)	15.1 (7.9)
Estancia en hospital	8.5 (5.5)	21.9 (8.6)	13.5 (9.8)	14.5 (5.6)	13.5 (9.8)
Estancia en UTI	3.3 (3.4)	2.5 (0.7)	3.8 (2.7)	7.0 (4.3)	3.8 (2.7)
Mortalidad (%)	3.38	13.3	26.5	17.9	9.1

n= número de pacientes, HAL= Hospital Ángeles de las Lomas, CMNLR= Centro Médico Nacional "La Raza", INCan= Instituto Nacional de Cancerología, HUT= Hospital de Traumatología, ()= desviaciones estándar.

4. Considerar abordajes más agresivos en el manejo del enfermo.

Dentro de las conclusiones del grupo de trabajo sobre problemas relacionados con la sepsis de la Sociedad Europea de Medicina de Cuidado Crítico⁸ se sugiere que las variables ideales para evaluar la disfunción o falla orgánica deben reunir los siguientes requisitos:

- Objetividad.
- Factibilidad.
- Simplicidad.
- Confiabilidad.
- Obtención a intervalos regulares.
- Especificidad para el órgano en cuestión.
- Dimensionalidad.
- Independencia del tipo de enfermedad.
- Independencia de la terapéutica.

Las variables empleadas en la mayoría de las escalas que evalúan disfunción orgánica reúnen los requisitos ya mencionados y son comunes: creatinina, bilirrubina, escala de Glasgow, índice de oxigenación, sólo existiendo heterogeneidad en el sistema cardiovascular en el que se usan presiones arteriales, frecuencia cardíaca, índices como la frecuencia cardíaca ajustada a la presión, uso de aminas o acidosis, como variables que lo califican.

Asimismo, el desarrollo de las diversas escalas ha empleado diversos puntos de corte, definidos de manera arbitraria o estadística, para cada una de las variables ya mencionadas. Lo anterior ha condicionado el surgimiento de varias escalas las cuales tienen diversas desventajas al momento de su aplicación en nuestra población, a saber, se han vali-

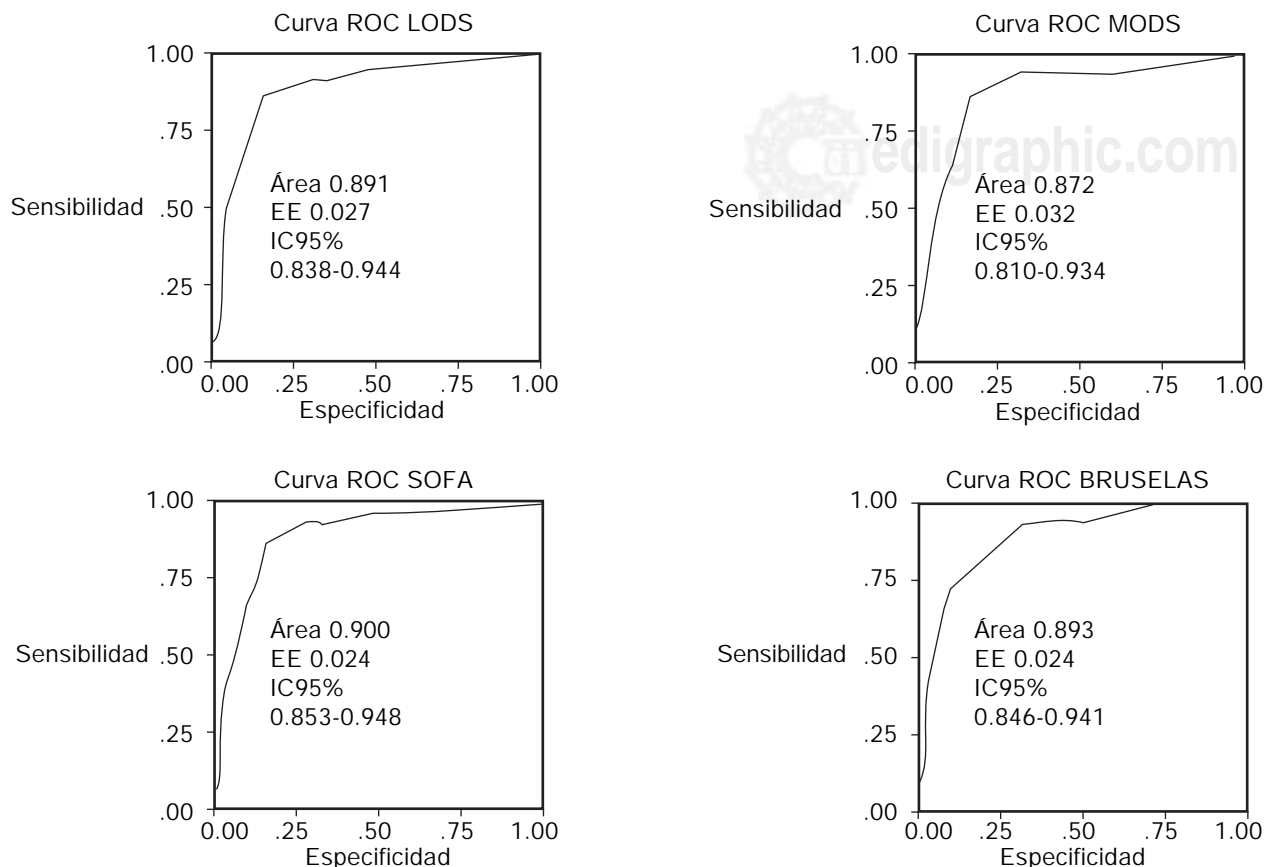
dado de manera retrospectiva, algunas se han diseñado en consenso, no en todos los estudios se realizó evaluación de su discriminación –capacidad del modelo de separar aquellos enfermos que predijo que sobrevivirían de aquellos enfermos que predijo que fallecerían–, y calibración –capacidad del modelo de describir el patrón de mortalidad en los datos–, se han aplicado en poblaciones anglosajonas, y, más importante aún, no hay estudios comparativos entre ellas.

El único estudio¹⁰ en el cual se compara una de las escalas, el MODS con el APACHE II y la escala propuesta por Knaus modificada en una población de 368 enfermos, se concluye que la escala MODS es útil en cuanto se emplee su delta, esto es, su valor máximo y su valor de ingreso a la UTI.

De ahí que se diseñara este estudio para llenar dos requisitos para su aplicación en nuestra población: validación en una muestra representativa de terapias intensivas mexicanas y su comparación en cuanto a capacidad discriminativa y calibración.

El primer objetivo se logró al reunir una muestra representativa de las patologías más frecuentes en nuestro país y aproximadamente en la misma proporción en que se presentan en una unidad de terapia intensiva polivalente mexicana, incluyendo enfermos oncológicos, cardiovasculares y traumatológicos.

Para el segundo objetivo en lugar de emplear sólo el cálculo estadístico de sensibilidad y especificidad, se aplicaron análisis de curvas ROC para calcular la capacidad discriminativa y se compararon a través de análisis que incluyó el coeficiente de correlación momento-producto de Pearson a dos colas, y también se empleó regresión logística múltiple para investigar la prueba de bondad de



Los segmentos diagonales son producidos por uniones.

Figura 1. Áreas bajo la curva ROC de las escalas de calificación de la gravedad del síndrome de disfunción orgánica múltiple.

Cuadro VI. Puntaje de las escalas de calificación de la gravedad del síndrome de disfunción orgánica múltiple.

Escala	Media ± DE	Intervalo	Mediana
Bruselas	2.4 + 2.5	0-14	2
SOFA	2.3 + 2.6	0-14	2
LODS	2.4 + 2.8	0-19	2
MODS	2.4 + 2.3	0-13	2

Cuadro VII. Áreas bajo la curva ROC, intervalos de confianza y valores de la χ^2 de Hosmer-Lemeshow de las escalas de calificación de la gravedad del síndrome de disfunción orgánica múltiple.

Escala	Área + EE	IC 95%	χ^2 H-L	p
Bruselas	89.3 + 2.4	84.6-94.1	6.23	0.182
SOFA	90.0 + 2.4	85.3-94.8	16.72	0.002
LODS	89.1 + 2.7	83.8-94.4	16.19	0.003
MODS	87.2 + 3.2	81.0-93.4	16.56	0.005

ajuste por la prueba de Hosmer-Lemeshow y determinar la calibración de cada modelo.

Con los anteriores datos se encontró que si bien todas las escalas tenían excelente capacidad discriminativa y no había diferencia estocástica entre ellas, pudiendo emplearse cualquiera de ellas con este fin, no ocurría lo mismo en cuanto a la calibración, siendo la única con significancia estocástica la escala Bruselas. De ahí que se proponga a la escala Bruselas como el sistema con buena discriminación y mejor calibración para evaluar el síndrome de disfunción orgánica múltiple.

BIBLIOGRAFÍA

1. Baue AE. Multiple, progressive, or sequential systems failure. A syndrome of the 1970. Arch Surg 1975; 110: 779-81.
2. Tran DD, Groeneveld ABJ, van der Meulen J et al. Age, chronic disease, sepsis, organ system failure, and mortality in a medical intensive care unit. Crit Care Med 1990; 18: 474-9.
3. Deitch EA. Multiple organ failure: Pathophysiology and potential future therapy. Ann Surg 1992; 216: 117-34.

4. American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: Definitions for sepsis and multiple organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med* 1992; 20: 864-74.
5. Knaus WA, Draper EA, Wagner DP et al. Prognosis in acute organ system failure. *Ann Surg* 1985; 202: 685-92.
6. Bernard GR, Doig G, Hudson LD, Lemeshow S, Marshall JC, Russell J et al. Quantification of organ failure for clinical trials and clinical practice. *Am J Respir Crit Care Med* 1995; 151: A323.
7. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: A reliable descriptor of a complex clinical outcome. *Crit Care Med* 1995; 23: 1638-52.
8. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med* 1996; 22: 707-10.
9. Le Gall JR, Klar J, Lemeshow S, Saulnier F, Alberti C, Artigas A et al. The logistic organ dysfunction system. A new way to assess organ dysfunction in the intensive care unit. *JAMA* 1996; 1276: 802-10.
10. Jacobs S, Zuleika M, Mphansa T. The multiple organ dysfunction score as a descriptor of patient outcome in septic shock compared with two other scoring systems. *Crit Care Med* 1999; 27: 741-4.

Correspondencia:

Dr. Luis David Sánchez Velázquez
Paseo de las Higueras 156-303,
Fracc. Paseos de Taxqueña.
Del. Coyoacán.
México, 04350, D.F.
Tel. 55-82-41-95.
Radiolocalizador. 52-30-30-30
Clave 14197.
E-mail: dbryan@prodigy.net.mx