

EL USO DE PROBABILIDADES SUBJETIVAS EN LA SIMULACION POR COMPUTADORAS DEL DIAGNOSTICO CLINICO¹

DRS. JOSÉ NEGRETE-MARTÍNEZ,^{2, 3} LADISLAO OLIVARES⁴ Y PEDRO SOLÍS-CÁMARA⁵

Se simuló el diagnóstico neurológico mediante el modelo de Ledley y Lusted para uso del teorema de Bayes. Una computadora programada para el objeto se alimentó con una matriz columna de 53 enfermedades y una matriz rectangular de 53 enfermedades por 295 síntomas y signos. Los elementos de estas matrices representan probabilidades subjetivas proporcionadas por neurólogos participantes en el trabajo. Se definió como diagnóstico de computadora aquella enfermedad, cuya probabilidad de ocurrir, dado el complejo sintomático, es la más alta. El sistema fue ensayado con 72 expedientes neurológicos, tomados al azar, lográndose 53 (73.6%) diagnósticos coincidentes con el clínico. Esto corresponde a un índice de correlación de 0.856 de un sistema binario, con valores teóricos separados entre 0.70 y 0.95 para límites de confianza de 95%. Se propone que la correlación encontrada es suficientemente alta como para considerar que se ha logrado una buena simulación del proceso mental del diagnóstico médico, considerado este como un proceso de identificación de formas en el sentido de la teoría de decisiones. Los resultados muestran además, que los estadísticos subjetivos son fuente de información adecuada para el diagnóstico por computadoras. Esto es de particular utilidad en aquellos campos de la Medicina donde no se cuenta con estadísticos objetivos utilizables. (GAC. MÉD. MÉX. 97: 1254, 1967).

UNO DE LOS procedimientos más exitosamente usados en el diagnóstico

médico por computadoras es el que hace uso del modelo probabilístico propuesto por Ledley y Lusted.¹ Consta éste de un universo patológico probabilístico finito, integrado por subconjuntos de entidades de naturaleza excluyentes: las enfermedades. Cada una de ellas, a su vez, se encuentra integrada por subconjuntos de síntomas, dis-

¹ Trabajo presentado en la sesión ordinaria del 5 de abril de 1967.

² Académico numerario.

³ Instituto de Estudios Médicos y Biológicos, Universidad Nacional Autónoma de México.

⁴ Hospital Nacional de Neurología, Secretaría de Salubridad y Asistencia.

⁵ Centro de Cálculo Electrónico, Universidad Nacional Autónoma de México.

tribuidos en el universo patológico como entidades no necesariamente excluyentes.

La aplicación del teorema de Bayes sobre este modelo ha generado experiencias de resultados alentadores, en el terreno del diagnóstico automático.²⁻⁶

Surge desde luego la pregunta de si existe o no alguna semejanza entre el proceso mental del diagnóstico y el proceso diseñado para las máquinas. Se ha postulado que en ambos casos se trata de un proceso de reconocimiento de formas y que el clínico utiliza primordialmente una información estadística subjetiva acumulada al través de su experiencia con enfermos.² Si este postulado es valedero, los estadísticos subjetivos del clínico podrán, adecuadamente extraídos y expresados, servir de base para el diagnóstico automático. Poner a prueba esta hipótesis es el propósito del presente trabajo.

Dada la multiplicidad de acepciones de algunos de los conceptos usados, hemos considerado pertinente el agregar un glosario de términos como apéndice del trabajo.

MATERIAL Y MÉTODOS

El diagnóstico según el modelo de Ledley y Lusted¹ consiste en una lista de enfermedades enunciadas en orden decreciente de probabilidad señalándose al margen la cifra de dicha probabilidad.

Esta probabilidad se calcula por aplicación del teorema de Bayes:

$$P(Y_i|X) = \frac{P(Y_i) \pi P(X/Y_i)}{\sum_{\text{Toda } X} (P(Y_i) \pi P(X/Y_i))}$$

i) En una forma directa, cuando X es el conjunto de síntomas que presenta el enfermo. Este conjunto pertenece al universo patológico: $(X_1, X_2, X_3, \dots, X_n)$, (Fórmula 1).

ii) O bien en variante, cuando X es el conjunto de síntomas que presenta el enfermo y que pertenecen al conjunto patológico $(X_1, \bar{X}_1, X_2, \bar{X}_2, X_n)$, (Fórmula 2). Véase glosario para la explicación de la notación.

En el presente trabajo se usó exclusivamente la primera fórmula dada la extrema dificultad que para el clínico representa el expresar probabilidades subjetivas sobre ausencia de síntomas.

Una variante usada en nuestros experimentos consistió en considerar que cualquier síntoma es capaz de presentarse en cada uno de los padecimientos, por improbable que su aparición se estime. De esta forma todos aquellos síntomas que a criterio de los clínicos "nunca se presentan" en determinada enfermedad, se consideraron con probabilidad 0.001. El haberles asignado una probabilidad cero habría sido equivalente a aceptar que se opera en un sistema "sin ruido" o sea uno a prueba de errores en la recolección e interpretación de síntomas.

Teniendo en cuenta que la computadora produce un espectro probabilístico diferencial, a diferencia del diagnóstico del clínico que es casi siempre la decisión de una sola enfermedad, fue necesario designar como "diagnóstico de computadora" a aquella enfermedad cuya probabilidad de ser la causa de los síntomas fuese la más alta en el espectro probabilístico producido por

la máquina. Esto implica un mecanismo de decisión obvio, impuesto al sistema simulador.

Se escogieron 52 enfermedades neurológicas, entre las más frecuentemente observadas en consulta externa de neurología, según estadística colectada por uno de los autores.⁷ Este grupo de enfermedades constituyó el universo patológico a reconocer por la computadora (Tabla 1).

diante cuestionario en los que se hacía elegir al clínico, entre dos enfermedades o dos de los síntomas de cada enfermedad, aquel que se considerase más frecuente. Lograda la ordenación, se asignaron valores porcentuales de frecuencia a cada una de las enfermedades y de los síntomas. En la realización de este procedimiento se aplicaron ciertas reglas: a) todas las enfermedades al igual que algunos síntomas se conside-

TABLA I

01	Epilepsia	28	Espondilosis cervical
02	Hemorragia cerebral	29	Síndrome de compresión radicular lumbar
03	Embolia cerebral	30	Distrofia muscular progresiva
04	Infarto cerebral	31	Amiotonía congénita
05	Síndrome craneohipertensivo	32	Hematoma subdural crónico
06	Migraña	33	Hemorragia subaracnoidea espontánea
07	Secuelas de encefalitis	34	Miotonía distrofica
08	Poliomielitis aguda	35	Hematoma epidural
09	Parkinsonismo	36	Miotonía congénita
10	Neuralgia del V par	37	Acromegalia
11	Meningitis viral	38	Miastenia gravis
12	Tabes dorsal	39	Tétanos
13	Parálisis general progresiva	40	Cefalea de tensión
14	Corea degenerativa	41	Demencia senil
15	Meningitis tuberculosa	42	Parálisis facial periférica
16	Meningitis piógena	43	Herpes zoster
17	Neuropatía diabética	44	Polineuritis saturnina
18	Esclerosis lateral amiotrófica	45	Corea reumática
19	Polineuritis infecciosa aguda	46	Siringomielia
20	Polineuritis alcohólica	47	Secuelas de poliomieltis
21	Polineuritis arsenical	48	Cranecostenosis
22	Polineuritis carencial	49	Mongolismo
23	Secuelas de infarto cerebral	50	Laberintitis y síndrome de Menière
24	Secuelas de trauma obstétrico	51	Hidrocefalia
25	Secuelas de contusión cerebral	52	Encefalitis viral
26	Siringobulbia	53	Parálisis de Erb
27	Compresión medular		

Se elaboró una lista de "síntomas" neurológicos para cada una de las enfermedades escogidas según criterio de los médicos participantes.

Hecha la lista de padecimientos y la de sus respectivos síntomas se procedió a ordenar sus elementos según frecuencia de aparición. Esto fue logrado me-

raron mutuamente excluyentes (la suma de sus porcentajes de frecuencia es igual a cien); b) se excluyeron deliberadamente ciertos datos. Ellos fueron algunos datos clínicos y de laboratorio que aunque muy valiosos no forman parte del estudio rutinario de los pacientes y por ello no pueden formar parte de la

experiencia probabilística subjetiva del médico.

Basándose en la información así obtenida se construyó una matriz columnar y otra rectangular conteniendo, la primera, las probabilidades a priori de cada una de las 53 enfermedades neurológicas consideradas en tanto que la segunda señalaba, en cada elemento, la probabilidad condicionada del síntoma dada la enfermedad.

Un número de casos individuales fueron diagnosticados con la aplicación del teorema de Bayes basándose en la información de estas matrices.

Del archivo clínico del Hospital Nacional de Neurología se tomaron 72

casos para juzgar la simulación programada. Puesto que este servicio contiene, además de expedientes neurológicos, un número de expedientes psiquiátricos, los casos fueron seleccionados utilizando sólo los primeros.

La revisión de expedientes se hizo por lo menos dos veces en cada caso con objeto de evitar hasta donde fuera posible las omisiones así como los errores de interpretación. La información de cada expediente fue transcrita en machotes *ad hoc* que contienen una lista de 295 síntomas de la matriz rectangular, con su correspondiente número de codificación (Tabla 2). En el encabezado de cada uno se anotó el nú-

TABLA 2

Nombre del médico	Institución
Diagnóstico más probable	Diagnóstico de máquina
Otros diagnósticos probables	no llenar
.....
Nombre del enfermo	Expediente No. Edad
Sexo: 279 masculino	280 femenino
<i>Antecedentes y/o padecimientos coexistentes</i>	<i>Características generales del padecimiento</i>
210 Alcohólicos	Edad en que principia
131 De cardiopatía	274 Lactancia (menor de un año)
195 Familiares positivos	275 Niñez (1-13)
205 Luéticos	276 Adolescencia (14-21)
268 Eritroblastosis fetal	277 Madurez (22-50)
254 Exposición a plomo y/o arsénico	278 Vejez (mayor de 50 años)
132 Fiebre reumática	Principio:
070 Parto distócico	282 Brusco con mejoría progresiva
203 Signo de Combe	303 Gradual mejoría progresiva
063 Trauma craneano	283 Insidioso curso ascendente
062 Trauma raquídeo	Evolución:
144 Diabetes	285 Por brotes con remisión total
064 Hipertensión arterial	284 Por brotes sin remisión total
124 Insuficiencia hepática	281 Curso no progresivo
133 Insuficiencia respiratoria crónica	Duración:
134 Trombosis múltiples	293 0.1 mes
	294 1.3 meses
	295 Mayor de 3 meses

Síntomas encefálicos

194	Ansiedad o angustia	Convulsiones:
246	Incapacidad para la concentración mental	007 Focales
158	Depresión	006 Generalizadas
052	Insomnio	238 Ensoñaciones
193	Irritabilidad	072 Errores de la conducta
162	Agitación	071 Falla en la autocrítica
Alucinaciones:		005 Pérdida de la conciencia
059	Auditivas	053 Pérdida global de la memoria
237	Gustatorias	054 Pérdida localizada de la memoria
236	Olfatorias	069 Pérdida de la memoria reciente
058	Visuales	010 Mioclonias
009	Ausencias	003 Náusea
008	Automatismo	165 "Período lúcido"
060	Cambios de personalidad	073 Retraso psicomotor
001	Cefalea	172 Somnolencia
002	Hemicránea	004 Vómito
066	Confusión mental	
297	Negativismo	

Signos encefálicos

194	Ansiedad o angustia	007	Convulsiones focales
246	Incapacidad para la concentración mental	006	Convulsiones generalizadas
158	Depresión	160	Delirio
247	Euforia	123	Desorientación derecha-izquierda
193	Irritabilidad	155	Desorientación temporoespacial
162	Agitación	048	Disfasia o afasia cognositiva
056	Agnosia visual	049	Disfasia o afasia motora
059	Alucinaciones auditivas	072	Errores de la conducta
058	Alucinaciones visuales	067	Estupor
057	Apraxia	071	Falla en la autocrítica.
009	Ausencias	055	Labilidad emocional
008	Automatismo	073	C. I. <100
068	Coma	172	Somnolencia
066	Confusión mental	004	Vómito.

Síntomas de nervios craneales

	II		V
183	Amaurosis bilateral	011	Dolor facial tórpido
185	Amaurosis unilateral		VII
270	Estrabismo convergente	082	Hemiparálisis o hemiparesia facial
182	Fotofobia		VIII
184	Fosfenos	045	Acúfenos
040	Escotoma	044	Sordera o hipoacusia bilateral
041	Hemianopsia homónima derecha	099	Sordera o hipoacusia derecha
036	Hemianopsia homónima izquierda	100	Sordera o hipoacusia izquierda
	III	046	Vértigo o mareo
253	Crisis oculógiras		IX
042	Diplopia	050	Disfagia
043	Ptosis		

Signos de nervios craneales

	I	191	Pérdida no dióptrica de la agudeza visual bilateral
296	Anosmia bilateral	183	Amaurosis bilateral
037	Anosmia unilateral	185	Amaurosis unilateral
	II	149	Atrofia de la papila
207	Pérdida no dióptrica de la agudeza visual unilateral	098	Campos visuales tubulares

- 040 Escotoma
 096 Hemianopsia heterónima binasal
 097 Hemianopsia heterónima bitemporal
 041 Hemianopsia homónima derecha
 036 Hemianopsia homónima izquierda
 261 Hemorragia subhialoidea
 115 Papiledema
 286 Defecto de campo visual no incluíble en los anteriores
 156 Ausencia de reflejo pupilar a la luz
 157 Ausencia de reflejo pupilar a la acomodación
- III
- 287 Midriasis bilateral
 154 Midriasis unilateral
 288 Miosis bilateral
 153 Miosis unilateral
 122 Oftalmoplejia con midriasis unilateral
 289 Oftalmoplejia con midriasis bilateral
 043 Ptosis
 114 Pupila irregular
 159 Pérdida del reflejo consensual
 156 Ausencia del reflejo pupilar a la luz
 157 Ausencia del reflejo pupilar de acomodación
 240 Signo de Parinaud
- IV
- 151 Diplopia sin oftalmoplejia aparente
- V
- 117 Pérdida del reflejo corneano
 089 Zona gatillo facial
- VI
- 152 Parálisis del abductor del ojo
- VII
- 038 Agusia del tercio anterior de la lengua
 082 Hemiparálisis o hemiparesia facial
 186 Parálisis o paresia facial bilateral
- VIII
- 119 Hiperacusia unilateral
 044 Sordera o hipoacusia bilateral
 099 Sordera o hipoacusia derecha
 100 Sordera o hipoacusia izquierda
 046 Vértigo o mareo
- IX-X
- 035 Agusia de los dos tercios posteriores de la lengua
 163 Ausencia de reflejo nauseoso
 039 Voz nasal
 252 Afonía
- XII
- 118 Atrofia lingual
 047 Disartria
 121 Hemiparesia lingual
 290 Hemiatrofia lingual
- Signos no clasificados
- 178 Exoftalmos bilateral
 180 Exoftalmos unilateral
 199 Voz monótona
 304 Nafzinger

Signos motores

- 235 Parálisis o paresia de miembros superiores
 227 Cuadruplejia o cuadriparesia flácida
 228 Fatigabilidad vespertina
 168 Bradicinesia
 271 Movimientos involuntarios
- 177 Contractura muscular
 094 Movimientos espasmódicos
 024 Torpeza de movimientos
 127 Pérdida del control de esfínteres
 095 Tics

Alteraciones de la fuerza

- 078 Cuadruplejia o cuadriparesia espástica
 227 Cuadruplejia o cuadriparesia flácida
 229 Disminución del volumen de la voz
 208 Disnea por paresia respiratoria
 076 Hemiplejia o hemiparesia espástica
 080 Hemiplejia o hemiparesia flácida
 291 Hemiatrofia fasciocorporal
 079 Monoparesia espástica
 022 Monoparesia flácida
 256 Parálisis del ciático
 077 Paraplejia o paraparesia espática
 075 Parálisis del mediano
- 242 Parálisis o paresia de grupos musculares
 235 Parálisis o paresia de miembros
 301 Parálisis del ciático popliteo externo
 302 Parálisis cubital
 257 Parálisis crural
 258 Parálisis radial
 081 Paraplejia o paraparesia flácida
 217 Defecto motor no incluíble en los anteriores
 267 Debilidad muscular

Alteración del tono

- 177 Contractura muscular
 176 Hipotonía muscular generalizada
 245 Hipotonía vesical
- 248 Rigidez de músculos espinales
 218 Trismus
 201 Asimetría de reflejos

Reflejos osteotendinosos

170	Clonus aquiliano	125	Hiperreflexia localizada
167	Clonus rotuliano	101	Hiporreflexia o arreflexia de los cuatro miembros
190	Hiperreflexia de miembros inferiores	086	Hiperreflexia o arreflexia distal
074	Hiperreflexia de un miembro inferior	085	Hiporreflexia o arreflexia generalizada
120	Hiperreflexia de los cuatro miembros	102	Hiperreflexia o arreflexia localizada
084	Hiperreflexia generalizada		

Signos motores misceláneos

171	Adiadococinecia	230	Fenómeno miotónico
265	Ataxia de miembros inferiores	241	Flexibilidad cérica
104	Ataxia del tronco	027	Incoordinación unilateral
088	Ataxia generalizada	223	Laterpulsión
249	Atrofia de una extremidad	143	Nistagmus vertical
030	Atrofia muscular distal	148	Nistagmus horizontal
028	Atrofia muscular local	150	Nistagmus roto
029	Atrofia muscular proximal	181	Opistotonos
298	Atrofia muscular generalizada	127	Pérdida del control de esfínteres
087	Babinski bilateral	083	Pseudohipertrofia muscular
126	Babinsky unilateral	239	Reflejo de Moro asimétrico
161	Ausencia del balanceo de brazos	219	Risa sardónica
023	Mutismo acinético	129	Signo de la rueda dentada
168	Bradicinecia	103	Romberg
174	Chvostek	173	Temblores aleteante
106	Dismetría bilateral	108	Temblores de acción
105	Dismetría unilateral	034	Temblores finos
220	Espasmo laríngeo	107	Temblores de reposo
221	Espasmo muscular generalizado	095	Tics
222	Espasmo muscular localizado	251	Tortícolis
031	Fasciculaciones	128	Facies inexpressiva
	Movimientos anormales		Marchas anormales
110	Atetósicos	025	Distónicos
111	Balísticos	094	Espasmódicos
026	Balísticos	272	Involuntarios
109	Coréicos		Inclasificables
		139	Cerebelosa
		140	Parkinsoniana
		138	Steppage
		141	Tabética
		306	En tijera

Síntomas sensitivos

214	Ardor de la planta de los pies	187	Dolor lancinante de miembros inferiores
136	Calambres	015	Dolor lumbar
292	Dolor talámico	212	Dolor palmar
188	Crisis dolorosas viscerales	211	Dolor suboccipital
012	Dolor cervical	264	Hiperestesia cutánea
215	Dolor ciático	091	Pérdida de sensibilidad en dermatoma
213	Dolor del hombro	092	Pérdida de sensibilidad en la palma de la mano
016	Dolor epigástrico paroxístico	018	Pérdida de sensibilidad y/o parestesia distal global en miembros inferiores
200	Dolor facial paroxístico	019	Pérdida de sensibilidad y/o parestesia de la cara externa del muslo
017	Dolor generalizado en miembro inferior	299	Pérdida de la sensibilidad y/o parestesia global distal de los cuatro miembros
013	Dolor generalizado en miembro superior		
014	Dolor intercostal		
216	Dolor intermetatarsico		

Signos sensitivos

093	Disociación de la sensibilidad	264	Hiperestesia cutánea
209	Dolor en el trayecto de nervios periféricos	090	Nivel sensorial
273	Dolor a la compresión de metatarsianos	091	Pérdida de sensibilidad en dermatoma
206	Dolor muscular a la presión	092	Pérdida de la sensibilidad en la palma de la mano
137	Dolor a la percusión de apófisis espinosas	018	Pérdida de la sensibilidad y/o parestesia distal global en miembros inf.

019	Pérdida de la sensibilidad y/o parestesia de la cara externa del muslo	113	Rigidez de la nuca
020	Pérdida del sentido de vibración y de posición de los miembros inferiores	116	Signo de Kernig
262	Pérdida de la sensibilidad al dolor y a la temperatura en la cara	032	Signo de Lasague bilateral
243	Pérdida de la sensibilidad al dolor profundo	112	Signo de Lasague unilateral
		196	Signos de Brudzinski
		175	Trousseau

Síntomas no neurológicos

142	Disminución de la libido	225	Deficiencia en el desarrollo somático
198	Crecimiento rápido de la cabeza	255	Diarrea
226	Crecimiento anormal de pies y manos	135	Lipotimias
189	Impotencia	166	Constipación

Signos no neurológicos

164	Asimetría craneana	263	Estridor laríngeo
232	Atrofia gonadal	224	Facies mongólica
179	Aumento del perímetro cefálico	300	Gigantismo
233	Calvicie	145	Hipertermia localizada
231	Catarata	244	Hipertensión arterial postural
226	Crecimiento anormal de pies y manos	260	Mancha oscura gingival
225	Deficiencia en el desarrollo somático	147	Microcefalia
169	Deformación simétrica del cráneo	234	Motilidad limitada de columna cervical
250	Escoliosis	130	Trastornos tróficos de la piel
202	Rodilla de Charcot	266	Vesículas en el trayecto de raíz espinal
232	Atrofia gonadal		

Síntomas generales

065	Fiebre	192	Apatía
033	Malestar general	021	Astenia
061	Pérdida de peso corporal	051	Aumento de peso corporal
204	Anorexia		

Exámenes de laboratorio

197	LCR anormal	259	Anemia
-----	-------------	-----	--------

Notas aclaratorias del médico:

.....

mero de serie, el número de expediente y el último diagnóstico señalado en este.

La información así codificada en los machotes fue pasada a tarjetas perforadas para uso de una computadora tipo CDC G20. El programa permite procesar los casos en forma individual o en serie, sin límite de número, produciendo cada diagnóstico en tiempo aproximado de 18 segundos.

El listado producido por la máquina, para cada caso clínico, consiste en: a) el número de clave del diagnóstico

clínico en caso de que éste exista; b) una lista de números correspondientes al cifrado de los síntomas que presentaba el enfermo y finalmente, c) en seis renglones consecutivos, las seis enfermedades más probables. En cada renglón se lee: el número de clave, el nombre en español de la enfermedad y la probabilidad correspondiente de ser la causa del complejo sintomático del enfermo. Esta lista de enfermedades está arreglada por la máquina en orden decreciente de probabilidades (Tabla 3).

TABLA 3
DATO LEIDO ENFERMEDAD (CASO NUM. 8)

CON SINTOMAS	1	66	193	280	276	285	295	6	5	0	0	0	0	0	0	0	0	
Prob. de de la Enf.	1	Epilepsia																= 0.745694189
Prob. de de la Enf.	5	Síndrome craneohipertensivo																= 0.133842547
Prob. de de la Enf.	25	Secuelas de contusión cerebral																= 0.093689783
Prob. de de la Enf.	15	Meningitis tuberculosa																= 0.015296291
Prob. de de la Enf.	33	Hemorragia subaracnoidea espontánea																= 0.005736109
Prob. de de la Enf.	3	Embolia cerebral																= 0.005736109

RESULTADOS

Las matrices subjetivas que se utilizaron en este trabajo se han comprimido para su representación de la siguiente manera: la matriz columna está escrita por quintetas de números, una a continuación de la otra. En éstas, los dos primeros números representan el número de clave de la enfermedad (véase lista de enfermedades Tabla 1) y los tres últimos, la probabilidad a priori de la ocurrencia de dicha enfermedad multiplicada por mil. La matriz rectangular está representada por septetas de números escritas una a continuación de la otra. Los primeros dos números de cada una representan la clave de la enfermedad, los tres si-

guientes la clave del síntoma (véase clave de los síntomas en Tabla 2) y los dos últimos la probabilidad condicionada multiplicada por cien de que se presente el síntoma dada la enfermedad. Todos aquellos síntomas que no aparecen con una probabilidad asignada tienen la probabilidad 0.001, por programa.

El programa, en lenguaje FORTRAN, está disponible en la programación del Centro de Cálculo Electrónico de la UNAM.

La Tabla 4, presenta el número y porcentaje de coincidencias entre el diagnóstico y el de computadora para primeras, segundas, terceras, etc., probabilidades del listado de diagnóstico diferencial de la máquina. El porcen-

TABLA 4

<i>Orden de probabilidad decreciente del diagnóstico diferencial de la computadora</i>	<i>Núm. de casos coincidentes con el diagnóstico médico</i>	<i>Porcentaje redondeado de coincidencias</i>
Primera	53	73.6
Segunda	10	13.8
Tercera	2	2.7
Cuarta	1	1.4
Quinta	1	1.4
Sexta	2	2.8
Superior a la sexta	3	4.2
T o t a l	72	100.0

taje de coincidencias (73.6) que encajeza la lista en esta tabla, constituye el resultado de la simulación según el modelo propuesto por nosotros.

La Figura 1, que podríamos lla-

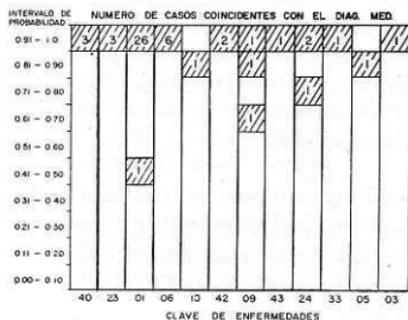


FIGURA 1

mar gráfica de certidumbre del diagnóstico, representa el valor probabilístico de la enfermedad, que constituye el diagnóstico de la máquina, en los

casos coincidentes con el diagnóstico médico.

Finalmente, la Tabla 5, contiene el número de veces que se diagnóstico clínicamente una enfermedad de la lista y su incidencia porcentual, así como la incidencia estimada en la matriz columna subjetiva. Nótese la buena correlación entre ambas informaciones (coeficiente de correlación 0.94).

DISCUSIÓN

Un método para la valoración de un sistema de simulación, consiste en alimentar los sistemas simulado y simulador, con la misma información y juzgar de la bondad de la simulación correlacionando lo obtenido a la salida de ambos sistemas. (Fig. 2, esquema superior).

Un experimento ideal sería el alimentar la mente del médico y la computadora con las mismas matrices, e

TABLA 5

Clave	Nombre de la enfermedad	No. de casos vistos Diag. Med.	Porcentaje redondeado del N° de casos vistos	Porcentaje de matriz columna	Porcentaje de mat. colum. normalizado y redondeado	Coefficiente de correlación Col. 2 y 4
40	Cefálea de tensión	4	5.6	8.0	10.3	0.94
29	Sin. comp. rad. lumb.	1	1.4	4.0	5.1	
23	Sec. inf. cerebral	5	7.0	8.0	10.3	
30	Dist. musc. prog.	1	1.4	0.5	0.6	
01	Epilepsia	32	44.4	25.0	32.3	
04	Infarto cerebral	3	4.2	2.5	3.2	
15	Meningitis tuberculosa	1	1.4	0.3	0.4	
06	Migraña	8	11.1	8.0	10.3	
10	Neuralgia del V par	2	2.8	0.1	0.1	
42	Parálisis facial perif.	2	2.8	8.0	10.3	
09	Parkinsonismo	4	5.6	4.0	5.2	
12	Tabes dorsal	1	1.4	0.1	0.1	
45	Corea reumática	1	1.4	4.0	5.2	
24	Secuelas traum. obst.	4	5.6	3.0	4.0	
33	Hemorragia subarac.	1	1.4	0.5	0.6	
05	Síndrome craneohipert.	1	1.4	1.0	1.2	
03	Embolia cerebral	1	1.4	0.5	0.6	
Total		72	100	77.5	100	

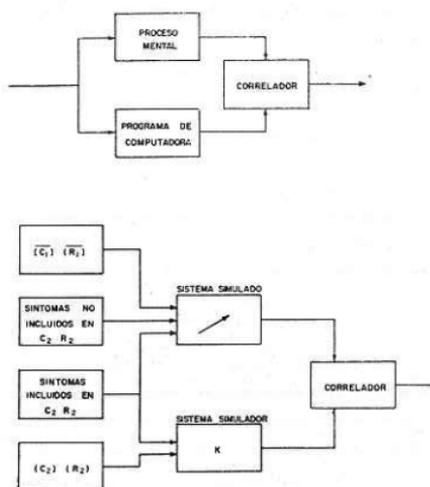


FIGURA 2

idéntico caso clínico comparando la "decisión diagnóstica" en ambos casos. Obviamente este no es nuestro caso ya que: *i)* El proceso mental que diagnóstica no es constante, ya que los diagnósticos médicos, se hacen por distintos individuos, y en ocasiones por la intervención simultánea de varios de ellos en un solo caso. Esta situación la hemos representado en el esquema de la parte inferior de la Figura 2, por un operador variable, (flecha inscrita en el rectángulo representativo); *ii)* a cada médico o grupo de médicos, que intervienen en el diagnóstico, se les debe asociar un par de matrices, que representan su "experiencia acumulada",⁸ en principio diferente de aquel par de ellas que alimentan al sistema simulador (esto se ha representado en la figura por la simbología de matrices, pero testando dichas matrices); *iii)* el proceso mental recibe una información

adicional, que la computadora no recibe, constituida por todos aquellos síntomas no incluidos en la matriz correspondiente que maneja la computadora; *iv)* en contraste, la computadora se puede considerar como un operador constante, (fórmula de Bayes, representada por K en la figura 2), que recibe una información matricial y una información sintomática reducida del caso clínico por analizar.

Las consideraciones anteriores nos hacen concluir que se está simulando un sistema probabilístico, en lo referente a operador e información, por medio de un sistema constante. Los resultados que pueden esperarse de la comparación, son pues de naturaleza estadística.

Si se toma el porcentaje de coincidencias entre ambos diagnósticos como una medida de correlación, se puede demostrar que considerando el sistema estadístico como un "sistema binario", esto es, asignándole un valor unitario a los casos en que la máquina coincide con el diagnóstico del médico y nulo cuando no coincide, la raíz cuadrada de la razón del número de coincidencias y número de enfermos, o la raíz cuadrada del porcentaje de coincidencias dividido entre 10, es el coeficiente e correlación de los datos (véase glosario).

Según el criterio anterior, 73.6% de coincidencias encontradas corresponde a un coeficiente de correlación de 0.856, con un valor teórico esperado (para límites de confianza de 95%) que oscila entre 0.70 y 0.95.

Este coeficiente de correlación indica que la simulación lograda es suficientemente buena como para compensar la desventaja de información con la

que trabaja la computadora o bien que la simulación es adecuada y la diferencia de información entre ambos sistemas no es muy importante.

El grado de certidumbre del diagnóstico de la computadora según se puede ver en la Figura 1 es muy alto, a tal extremo que nos permite calificar la decisión diagnóstica de enfática, pues en un solo caso se presentó una decisión diagnóstica con probabilidad inferior a sesenta y un centésimos.

Como puede verse en la Tabla 5, los enfermos muestreados, según se explicó en métodos, presentaron sólo 17 variedades de enfermedad de la lista de las 53 elegidas originalmente, como debía esperarse de las frecuencias *a priori* en la matriz columna. El éxito de la simulación, depende básicamente de esta circunstancia, ya que el contenido de información en cuanto a la probabilidad se refiere, es insignificante para aquellos síntomas o enfermedades infrecuentes.⁹

Una falla del modelo de Ledley y Lusted¹ es el considerar excluyentes a las enfermedades de un universo patológico. Sabemos que esto no es estrictamente cierto, pero para nuestros propósitos resulta una buena aproximación, ya que la coexistencia de dos enfermedades neurológicas es suficientemente rara como para no podersele asignar una frecuencia subjetiva.

Hasta ahora ha habido, en el diagnóstico por computadoras, una clara tendencia a la utilización de síntomas "objetivos", es decir, aquellos que son susceptibles de evaluación idéntica por distintos observadores. Esto ha obedecido

al deseo de minimizar el ruido en la información. Así en la mayor parte de las investigaciones en las que se emplea el teorema de Bayes, los autores han elegido especialidades que propician el uso de tales datos.

Puesto que el diagnóstico en medicina depende fundamentalmente de datos subjetivos, el empleo de ellos se hará tarde o temprano necesario en el diagnóstico automático.

Cuando aparecieron ideas para el diagnóstico por computadoras electrónicas, el servicio de estadística de los hospitales no estaba en posibilidad de suministrar información pertinente. El resultado fue el que muchos autores se vieron en la necesidad de recolectar sus propias estadísticas de la revisión de casos individuales.

Nuestro trabajo tiene, además de la intención original, de mostrar la simulación de un proceso mental por computadoras, el interés de mostrar que el uso de estadísticas subjetivas puede satisfactoriamente sustituir la información estadística objetiva en el diagnóstico por computadoras. Esta posibilidad probablemente elimine el serio obstáculo que para el diagnóstico automático constituye la falta de estadísticos objetivos adecuados.

No quisiéramos omitir de esta discusión un comentario acerca de lo inadecuado que el expediente clínico actual resulta como fuente de información, para el diagnóstico por computadoras. Baste mencionar lo impropio de su formato, la abundancia de sinónimas y la frecuente ambigüedad en la terminología. Estas circunstancias hacen que

la codificación necesariamente humana, sea lenta y poco fidedigna, lo que evidentemente es el cuello de botella para la aplicación del diagnóstico por computadoras en general.

Agradecimiento. Los autores agradecen al Sr. Miguel Soriano la elaboración del programa de computadora referido en este trabajo.

GLOSARIO

Universo patológico probabilístico finito

Conjunto de enfermedades del discurso en el que cada una de ellas lleva asociada su probabilidad *a priori* de ocurrir.

Notación de la fórmula de Bayes

Yi es la enfermedad cuyo número de clave es i, Xj es un síntoma presente cuya clave es el subíndice j, X̄j es un síntoma ausente, P () = probabilidad del evento especificado dentro del paréntesis. P (A/B), probabilidad de la ocurrencia del evento A condicionada a la ocurrencia de B. II significa un producto de n factores.

Sin ruido

Se refiere, por analogía al concepto de ingeniería de control, a un sistema en el que la información no está viciada por datos espurios.

Enfermedades

Se aplica el término de enfermedades a la denominación que usan los médicos para agrupar sus casos clínicos en los expedientes de hospital. En el entendimiento que esta denominación no siempre coincide con los conceptos corrientes de enfermedad.

Síntomas

Se aplica el término indistintamente a síntomas, signos, datos de laboratorio, gabinete, antecedentes, etc., en fin, a toda información que el enfermo aporta, pertinente a su padecimiento.

Matriz columnar

Arreglo de los números correspondientes a las probabilidades *a priori* de las enfermedades en forma de columna de números.

Matriz rectangular

Arreglo de los números correspondientes a

las probabilidades condicionadas de los síntomas en forma de 53 renglones por 295 columnas de números.

Decisión diagnóstica

El diagnóstico puede ser considerado tanto en el caso del médico como en el de la máquina, como una decisión.

Experiencia acumulada

Estadística acumulada por el médico en el curso de su práctica profesional.

Sistema binario (correlación de un)

En el cálculo del coeficiente de correlación para dos variables:

$$r = \frac{\sum a_i b_i}{\sqrt{(\sum a_i^2) \cdot (\sum b_i^2)}}$$

Podemos considerar el resultado del diagnóstico médico, como constante a igual a la usidad ($a_i = 1$) y b_i es diagnóstico de máquina, con valores $b_i = 0$ cuando no coincide con el diagnóstico médico y $b_i = 1$ cuando coincide. La ecuación (I) con estas restricciones se transforma en:

$$r = \sqrt{\text{No. de coincidencias} / \text{No. de casos}}$$

$$\text{o bien: } \sqrt{\% \text{ de coincidencias}} \div 10$$

$$r = \% \text{ de coincidencias} / 10$$

REFERENCIAS

- Ledley, R. S. y Lusted, L. B.: *Use of electronic computers to aid in medical diagnosis.* Proc. Inst. Radio Engineers 47: 1970, 1959.
- Conditional probability program for diagnosis of thyroid function. J.A.M.A. 183: 307, 1963.
- Overall, J. E. y Gorham, D. R.: *A pattern probability model for time classification of psychiatric patients.* Behavioral Sci. 8: 108, 1963.
- Warner, H. R., Toronto, A. F., Veasey, G. L. y Stephenson, R.: *A mathematical approach to medical diagnosis. Application.* J.A.M.A. 117: 177, 1961.
- Lodwick, G. S., Haun, C. L., Smith, W. E., Keller, R. F. y Robertson, E. D.: *Computer diagnosis of primary bone tumors. A preliminary report.* Radiol. 80: 273, 1963.

- Rinaldo, J. A., Scheinok, P. y Rupe, G. E.: *Symptom diagnosis. A mathematical analysis of epigastric pain*. Ann. Int. Med. 59: 145, 1963.
- Olivares, L. R.: *El paciente neurológico de consulta externa*. Rev. Med. ISSSTE 1: 311, 1964.
- Negrete-Martínez, J., Olivares, L. y Solís-Cámara, P.: *Valoración de la llamada experiencia acumulada*. Noveno Congreso Nacional de Ciencias Fisiológicas. Veracruz, 1966, p. 103.
- Takahashi, K.: Comunicación personal. 1965.

COMENTARIO OFICIAL

DR. CARLOS GUZMÁN FLORES¹

EN EL TRABAJO que nos acaban de presentar los Dres. Negrete, Olivares y Solís, referente al empleo de la estadística subjetiva del clínico para programar la computadora, los autores plantean la pregunta de si existe alguna semejanza entre el proceso mental del diagnóstico y el programa diseñado para la máquina. Es obvio que si la computadora es la máquina con la cual se trata de simular el diagnóstico clínico, su programa deberá de estar basado en un proceso lógico semejante en todo al que sigue el clínico para llegar a un diagnóstico. Sin embargo, una diferencia importante entre los dos procesos lógicos, es que en el del clínico la estimación de los síntomas con poca significación, en base a su experiencia, puede en un momento dado cambiar por completo para un caso particular. Este cambio en la valoración de un síntoma resulta de hecho de una asociación subconsciente difícil de analizar paso a paso y por lo tanto, difícil de reducir al lenguaje de la máquina. La computadora por otra parte, puede cambiar la estimación de los síntomas, pero nunca basada en un caso individual, sino en base a la estadística real que maneja.

El uso de la estadística de la experiencia individual se pone de manifiesto claramente a nivel de los médicos residentes en los

centros hospitalarios especializados. En estos centros se recibe un gran número de pacientes con enfermedades similares y en consecuencia, los residentes son capaces de realizar diagnósticos extraordinariamente elaborados, que a otro médico, con mayor experiencia que ellos, le sería muy difícil hacer en la práctica privada. Para aclarar este concepto diremos como ejemplo lo siguiente. En un hospital neurológico la mayoría de los casos de coma que se presentan, corresponden a accidentes vasculares cerebrales y por lo tanto, cuando en estos servicios el residente ve a un paciente comatoso, las probabilidades que tiene de acertar en el diagnóstico de accidente vascular son muy grandes. Por el contrario, en los centros hospitalarios de enfermedades endocrinas los comas diabéticos o insulínicos son los más frecuentes y para el residente de estos hospitales, la posibilidad de acertar el diagnóstico de diabetes es muy alta. Si ahora imaginamos un intercambio de residentes entre ambos hospitales y consideramos que cada uno de ellos pone en juego la estadística personal dada por su experiencia, encontraremos que el residente poco cuidadoso hará un diagnóstico equivocado en un alto porcentaje. Esto mismo ocurre a la computadora programada para el diagnóstico neurológico cuando es alimentada con historias clínicas de enfermos endocrinológicos y viceversa. Este hecho se debe a que el programa para el diagnóstico de una

¹ Académico numerario, Instituto de Estudios Médicos y Biológicos, Universidad Nacional Autónoma de México.

especialidad, tiene definidas las enfermedades en términos de la estadística con que se asocian, pero de ningún modo se encuentran definidas las enfermedades de otra especialidad.

Así pues, las consideraciones anteriores nos llevan a entender las limitaciones que tienen las computadoras en la realización del diagnóstico y sobre este mismo punto debemos señalar también, que dichas limitaciones son aún mayores que las del clínico que las programa y del cual se obtiene la estadística subjetiva que es posible reducir al lenguaje de las máquinas.

Cabe ahora preguntar si es posible alimentar a la computadora con la estadística subjetiva de varias generaciones de médicos y si el programa de la máquina puede ser mejorado a tal extremo que ésta resulte ser tanto o más precisa que el mejor médico. Teóricamente lo anterior es posible pero aquí nuevamente, debemos hacer énfasis en

el hecho de que para obtener la máquina perfecta es necesario que ésta sea alimentada por el médico con datos precisos de diagnóstico. En esta forma, la máquina irá siendo mejor a medida que la preparación del médico que le proporciona los datos también vaya siendo mejor.

El futuro del diagnóstico simulado por computadoras debe ser orientado hacia la enseñanza. El hecho de programar y convertir el proceso lógico del diagnóstico al lenguaje de la máquina, representa un buen ejercicio para el estudiante. Cuando se alimenta a la máquina con los datos recogidos en la elaboración de una historia clínica, se obtiene de inmediato el diagnóstico y éste puede ser cotejado con el obtenido por el estudiante. Así, éste podrá revisar paso a paso la recolección de sus datos así como la interpretación que de ellos haya hecho y así ir profundizando cada vez más en el estudio del proceso diagnóstico y su lógica.