

**SISTEMA ESPECIALISTA DE APOYO PARA EL DIAGNOSTICO Y  
SELECCION DE EXAMENES EN ENFERMEDADES GENETICAS**

Carlos Holzmann Ph.D.<sup>1</sup>, Ricardo Cruz Coke M.<sup>2</sup>, Mauricio Avaria S.

**ABSTRACT** Se desarrolla un sistema computacional de apoyo para el diagnóstico de enfermedades genéticas basado en información estadística. El sistema opera sobre la base de mejorar una hipótesis diagnóstica mediante la elección apropiada del próximo examen a realizar, para maximizar la reducción de la entropía informática o cantidad de incerteza de la hipótesis. El método se aplicó a un grupo reducido de trisomías con el fin de evaluar su comportamiento en pacientes reales. Se describen las principales características de la realización computacional en un computador personal, las que constituyen una interfaz amigable con el usuario médico.

### 1.- Introducción

En la actualidad existen diversos sistemas computacionales destinados a brindar apoyo al diagnóstico médico los que, de una u otra manera, intentan reproducir el razonamiento con incerteza que naturalmente realiza un especialista en pos de un diagnóstico definitivo [1-8].

El interés en desarrollar sistemas de apoyo al diagnóstico médico se justifica especialmente cuando se tratan áreas de la medicina en que el número de manifestaciones, síndromes y complicaciones superan las posibilidades de manejo habituales. Su aplicación resulta aún más justificada cuando el conocimiento es reducido y la información disponible está escasamente organizada, como es el caso de las enfermedades genéticas [9-14].

Para configurar el diagnóstico de las enfermedades que aquí interesan, dadas las características del conocimiento disponible que es básicamente estadístico, es indispensable la presencia del médico ya que éste debe aportar la porción no descrita (heurística). Ese conocimiento junto a la capacidad de manejo de información estadística por parte del computador, permite elaborar antecedentes propicios para alcanzar un diagnóstico confiable con el menor costo posible; costo, entendido en un sentido amplio, descrito más adelante.

### 2.- Representación del Modelo

Para enfrentar el problema del apoyo computacional al diagnóstico genético se requiere elaborar algún modelo lógico-matemático en el que se incorporen los conocimientos estadísticos adquiridos acerca de las enfermedades y sus correspondientes manifestaciones observables, como también la evaluación de las manifestaciones efectivamente observadas en el paciente. La figura 1. muestra una representación de las manifestaciones de un individuo, que facilita la explicación de lo que aquí se entiende por el problema del diagnóstico.

Manuscrito realizado bajo el auspicio del Departamento de Investigación y Bibliotecas, U. de Chile Proyecto DIB I-2275

(1) Dpto. Electricidad U. de Chile Casilla 412-3 Santiago Chile.

(2) Fac. Medicina U. de Chile Independencia 1027 Santiago Chile.

///Trabalho recebido em 30/06/87 e aceito em 05/08/87///

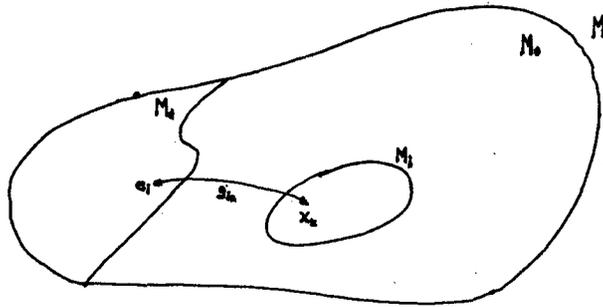


Figura 1.- Representación de las manifestaciones de un individuo, donde  $M$  es el universo de las manifestaciones,  $M_o$  es el conjunto de las manifestaciones observables (síntomas, signos, resultado de exámenes instrumentales),  $M_d$  es el conjunto de manifestaciones deducibles (síndromes, cuadros clínicos, enfermedades, etc),  $x_k$  es la  $k$ -ésima manifestación,  $e_i$  es la  $i$ -ésima enfermedad,  $g_{ik}$  es el grado de asociación o frecuencia de aparición de la manifestación  $x_k$  en la enfermedad  $e_i$  y  $M_i$  es el conjunto de las manifestaciones observables de la  $i$ -ésima enfermedad. Cada enfermedad induce un subconjunto  $M_i$ , contenido en el conjunto de las manifestaciones observables, que incluye a todas aquellas manifestaciones alguna vez descritas en pacientes con dicha enfermedad, de acuerdo al estado actual del conocimiento clínico.

## 2.1 El Problema del Diagnóstico

Llamaremos Hipótesis Diagnóstica, HD, al conjunto de todas aquellas enfermedades que explican las manifestaciones efectivamente observadas en un paciente, ordenadas de acuerdo a su importancia en la explicación.

Para perfeccionar la HD vigente y finalmente alcanzar un diagnóstico definitivo, es necesario el concurso de antecedentes adicionales a los encontrados hasta el momento de su formulación. Se requiere pues realizar una secuencia de exámenes que proporcionen la información necesaria para lograr una conclusión diagnóstica confiable. El orden en que se acopie esta información es importante pues de éste depende llegar más o menos eficientemente a un diagnóstico seguro y por consiguiente será también determinante en el costo del proceso.

La selección de los exámenes es una decisión compleja pues, además de la gran cantidad de ellos, deben considerarse los costos involucrados en su ejecución, a saber, el precio de los exámenes; la disponibilidad de los instrumentos, equipos o personal especialista; el riesgo y la condición física del paciente y el tiempo involucrado en su realización e interpretación. El análisis de la información que puede realizar un sistema computacional permite la búsqueda de alternativas con el objeto de minimizar los costos mencionados valorando, para ello, el contenido de información que proporciona cada examen. En el acto de la selección, el médico puede incorporar otros aspectos informáticos como son el conocimiento del medio hospitalario y su experiencia e intuición con el objeto de tratar de hacer aun más económico el proceso.

Se concluye entonces que sería ventajoso disponer de un criterio, acompañado de las herramientas de cálculo correspondientes, que permitan presentar al médico la secuencia de exámenes que, desde el punto de vista de su contenido informático, conduzca a un perfeccionamiento efectivo de la HD vigente hasta alcanzar un diagnóstico suficientemente confiable y permita adicionalmente una minimización del costo para los involucrados en la tarea diagnóstica.

## 2.2 Metodología

En esta sección se desarrolla el fundamento conceptual de la metodología utilizada para la selección de exámenes en base a información estadística; esto servirá al médico como un antecedente primario en sus decisiones. Especial atención merecen los conceptos de intensidad estadística de la enfermedad en un paciente y de entropía informática o cantidad de incerteza de la HD.

Cada una de las manifestaciones observadas en un paciente puede cuantificarse mediante su intensidad, es decir, la diferencia que se observa respecto de la condición normal. En este trabajo las intensidades se cuantifican normalizadamente en el intervalo  $[0,1]$ , correspondiendo el valor cero a la condición normal. Además, se utilizará el valor  $1/e = 0.37$  (\*) para la intensidad de las manifestaciones aun no observadas, el que se reemplazará por su real valor una vez realizado el examen correspondiente. De esta forma cada paciente define un vector de intensidades de sus manifestaciones,  $I$ , cuyas componentes tienen el valor  $1/e$ , en el caso de manifestaciones aun no observadas, o el valor de intensidad establecido a través del examen correspondiente.

A falta de información cuantificable más precisa, que relacione enfermedades y manifestaciones -como ocurre en otras áreas de la medicina- en este sistema de apoyo al diagnóstico de enfermedades genéticas se utiliza la frecuencia de aparición de la manifestación  $x_i$  en los pacientes cuyo diagnóstico ha incluido a la enfermedad  $e_i$ . A esta frecuencia se la llama grado de asociación  $g_{ik}$ . Un arreglo matricial de estos grados de asociación en filas, correspondientes a las enfermedades, y columnas a las manifestaciones, definen la matriz de asociación,  $G[g_{ik}]$ . Para efectos de este trabajo, los grados de asociación están normalizados en el intervalo  $[0,1]$ . Con el objeto de uniformar el valor total estadístico de cada enfermedad, adicionalmente se normalizan las filas de  $G$  respecto de la suma de sus entradas; de esa manera la matriz queda en la forma estocástica con la propiedad de que todas sus filas suman la unidad. Una explicación para la transformación realizada es que las enfermedades deben valorarse con igual certeza en presencia de todas sus manifestaciones.

Es conveniente aclarar que debido a que estas frecuencias de aparición generalmente han sido obtenidas en una población reducida, la información que contienen los grados de asociación no tienen validez estadística propiamente tal.

(\*) El valor  $1/e=0.37$  no es arbitrario; su sentido y conveniencia pueden comprenderse a la luz de las fórmulas de cálculo de la entropía informática de la HD en la sección 2.4 y en el capítulo de conclusiones.

Para discriminar acerca de la importancia relativa entre las distintas enfermedades es necesario contar con algún elemento de evaluación. Este debe ser un parámetro cuyo valor se relacione directamente con el conocimiento entre enfermedades y sus manifestaciones, así como con la condición conocida (observada) del paciente. En consecuencia se define el vector de intensidades estadísticas,  $U$ , de un paciente como

$$U = G[g_{iu}] * I. \quad (1)$$

La componente  $U_i$  del vector  $U$  representa la intensidad estadística con que se presenta la  $i$ -ésima enfermedad en el paciente. Este valor también puede interpretarse como el grado de pertenencia (estadística) de la enfermedad  $e_i$  al paciente.

Las intensidades estadísticas permiten cuantificar la importancia relativa de las distintas enfermedades en relación a las manifestaciones observadas, posibilitando el ordenamiento jerárquico de ellas, que fue sugerido en la definición de HD de la sección 2.1. En resumen y desde este punto de vista, la HD vigente sería el conjunto de todas las enfermedades (consideradas) ordenadas jerárquicamente de acuerdo a sus respectivas intensidades estadísticas.

### 2.3 Selección de Exámenes

La certeza, en el contexto del diagnóstico médico, sólo se consigue cuando se verifica la presencia de manifestaciones patognomónicas que, por naturaleza, son determinísticas. En el contexto de las enfermedades genéticas también existen dichas manifestaciones sin embargo, generalmente su constatación es costosa. Por ejemplo, el cariotipo es patognomónico para las trisomías pero son escasos los laboratorios que lo realizan y su valor es alto comparado con otras alternativas para establecer el diagnóstico.

El procedimiento médico para la selección de exámenes incluye normalmente información no descrita proveniente de la experiencia adquirida por el profesional, su intuición y el conocimiento del medio que lo rodea. En ocasiones esto podría inducir al usuario médico a alterar la jerarquización establecida en la HD, privilegiando una o varias enfermedades sobre las otras y realizar exámenes tendientes a confirmar esa suposición.

No obstante lo anterior la estrategia básica que se plantea en este trabajo para la selección de exámenes conducente al perfeccionamiento de la HD vigente, consiste en la valoración a priori de la información que proporcionaría la realización de cada uno de ellos. Esto permite un ordenamiento jerárquico de todos los exámenes posibles, en relación a su potencialidad de afectar a la actual HD. Un criterio objetivo que considera globalmente la HD vigente, sin privilegiar alguna enfermedad o grupo de ellas, lo constituye la evaluación de la sensibilidad de la entropía informática respecto de las manifestaciones observables, es decir, la razón de cambio en la cantidad de incerteza contenida en la HD vigente.

## 2.4 Entropía Informática

Por analogía con el caso de la termodinámica, una HD es equivalente a un macroestado, entendiéndose por éste a cualquier condición probabilística de un sistema. En un macroestado determinado, se pueden distinguir distintas valoraciones de los elementos constitutivos; cada una de éstas recibe el nombre de microestado. Para el caso del macroestado HD vigente se distinguen, por su intensidad estadística, las enfermedades que lo componen, constituyendo éstas los microestados correspondientes.

Desde el punto de vista de la teoría de la información, la probabilidad de ocurrencia de un cierto evento  $x_k$ , que también puede considerarse como la medida de su incerteza, denotada por  $p(x_k)$ , está relacionada con la cantidad de información  $i(x_k)$  que este evento aporta mediante :

$$i(x_k) = \ln(1/p(x_k)) \quad (2)$$

Entonces, el valor esperado de la información proporcionada por una distribución probabilística queda expresada por:

$$E(x) = E_0 \sum_k p(x_k) * \ln(1/p(x_k)) \quad (3)$$

La cantidad  $E$  lleva el nombre de entropía o cantidad de incerteza de la distribución y  $E_0$  es una constante de normalización.

En el caso considerado en este trabajo, la distribución probabilística está constituida por las intensidades estadísticas  $U_i$  del macroestado HD. La entropía  $E$  de cada microestado  $U_i$  está dada por

$$E_i = e * U_i * \ln(1/U_i) \quad (4)$$

donde el factor  $e$  normaliza el máximo de  $E_i$  al valor unitario. En consecuencia, la entropía  $E$  del macroestado HD será

$$E = (1/N) \sum_{i=1}^N E_i \quad (5)$$

lo cual constituye el valor promedio de las microentropías.

El valor máximo de la entropía de cualquier microestado se ha normalizado en el valor unitario, el cual representa la situación de mayor ignorancia respecto a la existencia de la correspondiente enfermedad en el paciente. La realización de exámenes incorpora información a la HD que modifica el valor de su entropía o cantidad de incerteza.

En consecuencia, el examen que más conviene realizar es aquel que tiene la potencialidad de producir mayor disminución en la cantidad de incerteza de la HD vigente. Esta evaluación "a priori" puede determinarse mediante la sensibilidad de la entropía respecto de la intensidad de cada una de las manifestaciones aun no evaluadas, siendo su expresión

$$S_k = \partial E / \partial I_k = (-e/N) * \sum_i [\ln(U_i) + 1] * g_{ik} \quad (6)$$

donde  $k$  es el índice de la  $k$ -ésima manifestación aun no evaluada.

En consideración a lo anterior, cada manifestación aun no evaluada lleva asociada un valor  $S_k$  y mientras menor sea tal valor mayor será la información que aportará la realización del examen correspondiente. En conclusión, los valores de las sensibilidades permiten jerarquizar los exámenes a realizar de acuerdo a la potencial variación (valor a priori) en la entropía de la HD vigente.

Una vez realizado el examen de mayor jerarquía informática, se conoce el valor de su intensidad,  $I_k$ , lo cual permite reevaluar la HD utilizando (1), mediante la incorporación del verdadero valor de la intensidad de la manifestación recién determinada en la k-ésima componente del vector I, en reemplazo del valor 0.37 supuesto inicialmente.

Lo anterior establece un criterio global y objetivo para cerrar y continuar el ciclo hipótesis diagnóstica-selección de exámenes. La selección del nuevo examen estará informada por este elemento de juicio basado en la disminución máxima de la cantidad de incerteza de la HD vigente. Obviamente este procedimiento, por su complejidad de cálculo, difícilmente podría practicarse en forma mental.

Es importante mencionar que la expresión de la entropía en (4) depende tanto del grado de asociación entre manifestaciones y enfermedades, que son cantidades fijas, como del estado del paciente,  $I_k$ , lo cual implica, a través de (6), que el orden jerárquico de los exámenes a realizar depende de dicho estado del paciente.

## 2.5 Truncamiento del Proceso de Aproximación al Diagnóstico.

El ciclo hipótesis diagnóstica-selección de exámenes conduce gradualmente a configurar una HD cada vez más informada y con el mínimo de exámenes. Sin embargo, se hace necesario disponer de algún parámetro que mida la magnitud de los antecedentes considerados y permita decidir el término del proceso antes de realizar todos los exámenes posibles. Con tal fin se define la cantidad de certeza con que se ha evaluado la intensidad estadística de la i-ésima enfermedad mediante

$$C_i = (1 - E_i) * 100 \quad [\%] \quad (7)$$

donde  $E_i$  es la entropía del microestado correspondiente a la i-ésima enfermedad dada en (4). Claramente, cuando todas las manifestaciones observables han sido observadas se verifica que  $C_i = 100\%$  para cada enfermedad, sin embargo la rapidez en el crecimiento de C es diferente para cada enfermedad poniendo de manifiesto anticipadamente la certeza del diagnóstico más probable.

## 3.- Resultados

A continuación se entregan algunos resultados obtenidos en la aplicación práctica de la metodología desarrollada. Para ello se han considerado los datos extraídos de fichas de pacientes del Servicio de Genética del Hospital Clínico José Joaquín Aguirre de la Universidad de Chile.

Se dispuso el sistema computacional para considerar las Trisomías 8, 9, 13, 18 y 21. La elección de estas anomalías se fundamenta en que poseen un conjunto muy grande de manifestaciones (>100) y extraordina-

riamente compartidas por ellas. Además, las trisomías presentan la ventaja adicional de que un cariograma del paciente basta para tener configurado un diagnóstico exacto. Claro está que el cariotipo no se incluyó entre las manifestaciones observables.

Se realizaron diagnósticos de ocho pacientes con algunas de las enfermedades consideradas arrojando resultados concordantes en todos los casos, después de examinar sólo dieciocho manifestaciones. A manera de ejemplo se muestra el conjunto de gráficos que ilustran la variación de la intensidad estadística, la cantidad de certeza y la entropía versus los exámenes realizados para un paciente de trisomía 21 (comprobado a través de un cariograma). Las fichas utilizadas en la comprobación del comportamiento del sistema incluyen un conjunto de dieciocho manifestaciones cuyas intensidades se indicaron a través de los signos -, +/-, +, ++, +++ que representan respectivamente las intensidades 0,50,33,66 y 100%. Debido a que no se anotan en la ficha aquellas manifestaciones examinadas pero no encontradas, la información que provee la ficha es incompleta y por ende la cantidad de certeza que se pueda alcanzar es relativamente baja en términos absolutos.

Es conveniente dejar en claro que en ausencia completa de observaciones de un paciente, la cantidad de incerteza en su HD es máxima; efectivamente, computando (1) las intensidades estadísticas de todas las enfermedades es  $U_i = 1/e$  (ya que el vector de intensidades  $I$  tiene todas sus entradas con el valor  $1/e$ ) y en consecuencia la cantidad de incerteza de cada enfermedad es  $E_i = 2.71 * U_i * \ln(1/U_i) = 1$  y la de la HD es  $E = (1/N) \sum E_i = 1$ .

La figura 2 muestra el efecto de la selección de exámenes en el valor de la intensidad estadística, en donde se observa que, con la información que incorporan los exámenes realizados, el crecimiento en el parámetro es sostenido sólo en el caso de la enfermedad que efectivamente posee el paciente; en cambio, en todas las demás, pasado un cierto número de exámenes, este índice detiene su crecimiento. Los exámenes realizados corresponden en su mayoría a medidas corporales y detección visual de malformaciones evidentes, tal como puede verificarse en la tabla 1, lo que implica un costo muy reducido en comparación con un cariograma.

Debe observarse que debido a que la ficha clínica habitualmente no incluye información acerca de las manifestaciones ausentes en el paciente, las curvas de la intensidad estadística resultaron monótonicamente crecientes, debiendo en realidad decrecer si esas observaciones se hubiesen anotado. Existe una diferencia conceptual (cuantificada) entre una manifestación observada como normal (intensidad 0.0) en relación con no haber observado dicha manifestación (intensidad  $1/e=0.37$ ).

En la figura 3 se han graficado las curvas correspondientes a las microentropías. Resulta evidente que la enfermedad que mayormente ha variado su entropía es la trisomía 21 indicando con ello que su cantidad de incerteza es menor luego de realizada la secuencia de exámenes señalada en el gráfico.

La figura 4 muestra el comportamiento de la cantidad de certeza para las distintas enfermedades consideradas. Todas ellas cambian con distinta rapidez, a medida que se incorporan los resultados de los

exámenes. Es necesario agregar que debido a que los valores de las intensidades estadísticas tienen un valor inicial de 0.37, el rango de variación del parámetro intensidad estadística se encuentra en [0.37,1.0] con lo cual el rango de variación del grado de certeza se restringe a un rango menor. Claramente la cantidad de certeza adquirida respecto de la enfermedad que se ha diagnosticado para el paciente, es superior a los valores alcanzados en las otras enfermedades, aunque no sea un valor suficientemente grande debido a las razones dadas anteriormente. En cuanto el usuario médico se dé por satisfecho de la certeza con que se establece la principal enfermedad de la HD, puede truncar el procedimiento para considerar como definitiva a la HD vigente.

La figura 5 permite comprobar que la secuencia generada por la metodología desarrollada efectivamente va disminuyendo la entropía de la HD a medida que se incorporan los resultados de los exámenes. La tendencia de la curva indica que un mayor número de exámenes permitiría alcanzar el valor final nulo en forma asintótica. Puede observarse además, que en el caso del examen 75, que resultó negativo, la entropía tiende a crecer denotando con ello la incerteza que introduce en la formulación de la HD original. La metodología predijo para el caso de esa manifestación, una variación negativa de la entropía mayor que para los restantes exámenes, sin embargo el examen resultó negativo cambiando con ello las condiciones supuestas en el cálculo.

En todos los gráficos se han codificado los exámenes de acuerdo a su posición en la base de conocimiento. La HD original se realizó con las dos primeras manifestaciones que aparecen en la tabla 1 y su perfeccionamiento, con la secuencia allí establecida, que es la resultante del proceso de selección desarrollado en este trabajo.

#### 4.- Desarrollo computacional

El desarrollo computacional se realizó en un microcomputador IBM-PC con 640 Kb. El lenguaje utilizado fue BASIC compilado y se utilizaron alrededor de 250 Kb en programas y bases de datos.

Una de las grandes causas del rechazo de muchos médicos a sistemas de apoyo computacional se encuentra en la dificultad de su uso. Para solucionar ese aspecto se dispuso una interfaz por software, para ofrecer una interacción "amigable" con el usuario. Esta interfaz se desarrolló en base a menús, pantallas en varios niveles con comandos para cada nivel, una enciclopedia con información acerca de las enfermedades y sus correspondientes manifestaciones, comparación automática de las medidas corporales con los estándares, etc. Además, se proveen estructuras ordenadas para la adquisición de la información que debe entregar el usuario, incluyendo la capacidad de modificar, de acuerdo a los diagnósticos realizados, la base de conocimiento, como también la posibilidad de emitir fichas clínicas para cada paciente.

Una sesión típica utilizando el sistema desarrollado comienza cuando el médico ingresa su palabra clave que lo habilita para acceder al sistema. Esta precaución es necesaria debido a que se manejan fichas clínicas confidenciales. A continuación se despliega el menú principal que incluye los siguientes elementos:

A : Enciclopedia de Enfermedades y Manifestaciones  
B : Anamnesis del Paciente  
C : Diagnóstico y Selección de Exámenes  
P : Impresión de Ficha Clínica  
X : Fin de la Sesión.

Al presionar cualquiera de las teclas indicadas se tiene acceso al siguiente nivel. La enciclopedia corresponde a una base de datos de alrededor de 120 Kbytes de memoria en que se condensa casi la totalidad del conocimiento escrito acerca de las enfermedades genéticas más comunes.

La anamnesis próxima y remota del paciente junto a un completo examen físico puede realizarse a través de la función bajo el rótulo B. Secuencialmente se solicitan antecedentes personales y familiares del paciente, luego en base a una división del cuerpo humano, bastante común en la práctica médica, se solicitan ciertas medidas corporales con las cuales se pueden configurar manifestaciones típicas de las enfermedades genéticas. Para facilitar aún más la labor del médico, se provee la comparación automática de los valores anteriores con los estándares de la población; ello permite valorar rápida y objetivamente las intensidades de las manifestaciones. Por último, todos los datos anteriores configuran una ficha clínica del paciente que se almacena en diskette. Un diskette 2S/2D puede llegar a contener unas 300 fichas de pacientes.

Con la función C se ingresa al programa de selección de exámenes propiamente tal, el cual recibe los datos ya ingresados por el usuario y estructura la HD inicial. El proceso cíclico que sigue a continuación, elabora la secuencia jerarquizada de exámenes que deberían realizarse de acuerdo a la metodología descrita. El médico usuario puede alterar el proceso, de acuerdo a información no escrita ni programada, forzando la secuencia de exámenes que considere más conveniente. Por ejemplo, puede no realizar el examen que aporta mayor cambio en la entropía informática por razones tales como su intuición educada acerca de la enfermedad más probable o sencillamente por considerar que el paciente no dispone de tiempo, dinero o no está capacitado físicamente para soportar un examen. Haciendo uso de esta capacidad el usuario puede comenzar el proceso por aquellas manifestaciones cuya evaluación puede realizar fácilmente en su consulta, para continuar con aquellas de más costosa ejecución si fuese necesario. Todo lo anterior a la luz de la jerarquización de exámenes establecida mediante la metodología aquí desarrollada.

## 5.- Conclusiones

A medida que aumente el conocimiento acerca de las enfermedades genéticas, podría incorporarse al sistema la capacidad de explicar los diagnósticos alcanzados. Asimismo, podría mejorarse la calidad de la información disponible permitiendo una estructuración mayor de las enfermedades y la constitución de una verdadera base de conocimiento. Desde esa perspectiva, el sistema de apoyo al diagnóstico presentado en este trabajo, se encuentra en vías de llegar a ser un sistema experto propiamente tal.

Una modificación operativa de gran conveniencia es el cambio del lenguaje de programación por uno de más rápida ejecución tal como el Turbo Pascal, lo que agilizaría la interacción con el usuario.

Resulta significativo el contar con una valoración de la ignorancia que se tiene inicialmente, acerca de la enfermedad del paciente. El valor  $1/e$  es ese valor ya que en él se produce el máximo de la entropía de cualquiera enfermedad, por lo tanto cualquier información que se incorpore a la HD por el primer examen determinará un decrecimiento de la entropía de la enfermedad correspondiente.

Es interesante agregar que la dependencia que posee la secuencia de exámenes respecto del estado del paciente, puede producir un ordenamiento distinto de los exámenes para cada paciente. Esto implica que el procedimiento desarrollado sería imposible de traducir en un protocolo uniforme igualmente aplicable a todos los pacientes.

Los resultados obtenidos indican que el procedimiento utilizado en la selección de exámenes, efectivamente consigue alcanzar un diagnóstico diferencial razonablemente confiable, basado en el seguimiento de una secuencia óptima para la minimización de la entropía de la HD. En otras palabras, la secuencia de exámenes realizada maximiza la cantidad de información que se aporta en cada paso del ciclo de perfeccionamiento de la HD.

El comportamiento del sistema adicionado a la gran capacidad de interacción con el usuario hacen esperar una buena acogida del sistema por parte de los médicos y la posibilidad de un uso masivo en policlínicos de áreas periféricas, en donde no siempre se cuenta con especialistas y/o equipos adecuados para la realización de exámenes muy complejos.

**TABLA 1**  
Exámenes de manifestaciones de Trisomía 21

Código	Manifestaciones
74	Retardo mental
34	Cuello Alado
75	Defecto cardíaco congénito
32	Implante bajo de orejas
18	Fascies Aplanada
35	Cuello corto
69	Hipotonicidad
52	Clinodactilia
92	Cresta simiana
7	Fisura párpado oblicua
8	Pliegue Epicanto
17	Puente Nasal Bajo
48	Dedos Cortos
62	Hiperextensibilidad
2	Occipucio Plano
46	Extremidades Cortas
47	Manos Chicas
30	Lengua Fisurada

## Bibliografía

- [1] Thomas Whalen, "Decision Making Under Uncertainty with Various Assumptions about Available Information" IEEE Transactions on Systems, Man and Cybernetics, Vol SMC-14 N°6 Nov/Dec 1984
- [2] J.A. Reggia, P.Y. Wang and D.S. Nau, "Minimal Set Covers as a Model for Diagnostic Problem Solving" Proceedings MEDCOMP'82 pp 340-347
- [3] W. Rogers, B. Ryack, G. Moeller, "Computer Aided Medical Diagnosis : Literature Review", Int. J. Biomed. Comput., Vol. 10, N° 4, Aug. 1979, pp 267-289
- [4] G.Soula, J.Gouvernet, A. Barre, J.L. San Marco, "Application of Fuzzy Relations to Medical Decision Making." Medinfo 80, Proc. of the Third World Conf. on Medical Informatics, Tokyo, Japan, 29 Sept.-4 Oct., 1980, pp.844-848
- [5] H. Welchsler, "A Fuzzy Approach to Medical Diagnosis" Int. J. Bio-Computing, Vol. 7, 1976, pp. 191-203
- [6] Daniel G. Bobrow, Sanjay Mittal, and Mark J. Stefik, "Expert Systems : Perils and Promise" Communications of the ACM September 1986, Vol.29 Number 9
- [7] Dana S. Nau, University of Maryland, "Expert Computer Systems" Computer Feb. 1983
- [8] Anne Fagot Largeault, Ph.D., M.D. Universidad de París, "La Simulación del Razonamiento Médico" Mundo Científico N°53 Vol. 5, pp 1236-45
- [9] H. M. Pashayan, M.D., "What else to look for in a child born with a cleft of the lip and/or palate", Cleft Palate Journal, Vol. 20, N°1, Jan 1983
- [10] D. W. Smith, Recognizable Patterns of Human Malformations, W. B. Saunders Company 3rd. Edition Philadelphia, 1982
- [11] Daniel Bergsma, M.D., M.P.H., Birth Defects Compendium 2ª Edition 1979
- [12] Instituto Nacional de Estadísticas Estadísticas sobre Egresos Hospitalarios por Grupos de Edad según Causas de Morbilidad Chile 1983
- [13] Richard M. Goodman, Robert J. Gorlin, Atlas of the Face in Genetics Disorders Second Edition
- [14] D. W. Smith, Atlas de Malformaciones Somáticas en el Niño Editorial Pediátrica.

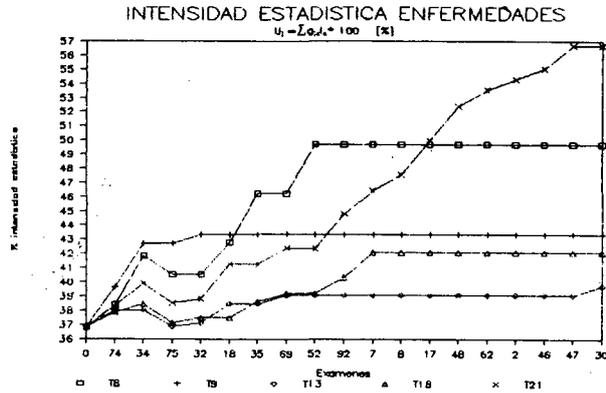


Figura 2. Gráfico de las intensidades estadísticas de las trisomías 8,9,13,18 y 21 para el caso de un paciente de trisomía 21.

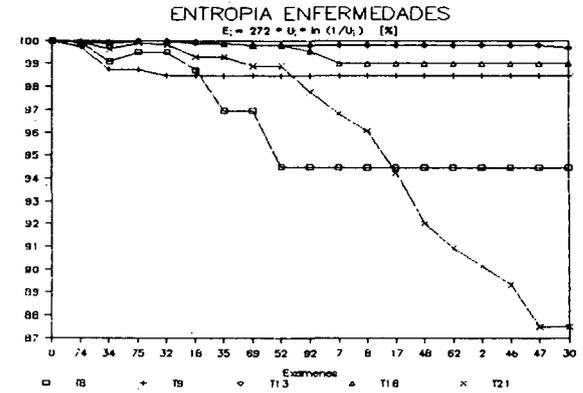


Figura 3. Gráfico de la entropía de cada una de las trisomías 8,9,13,18 y 21 para el caso de un paciente de trisomía 21.

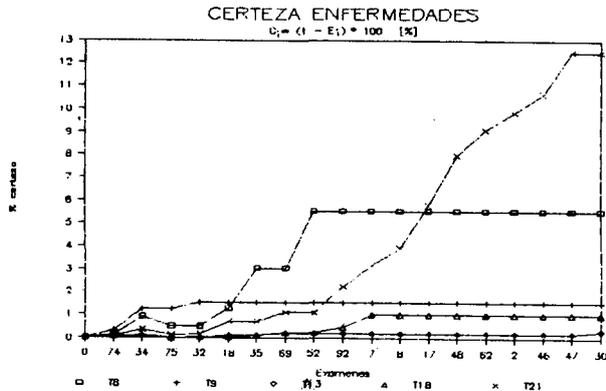


Figura 4. Gráfico de la cantidad de certeza de las trisomías 8,9,13,18 y 21 para el caso de un paciente de trisomía 21.

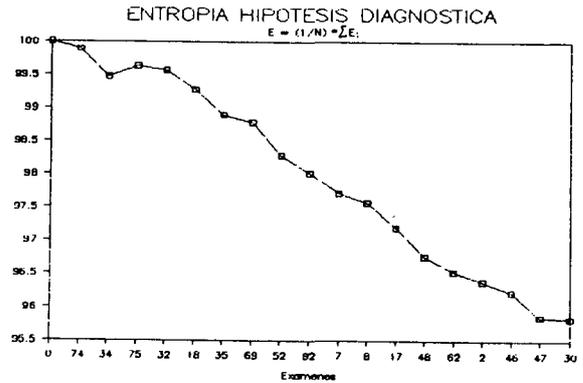


Figura 5. Gráfico de la entropía de la Hipótesis Diagnóstica para el caso de un paciente de trisomía 21.

SPECIALIST SYSTEM FOR GENETIC DISEASE DIAGNOSIS  
AND EXAMINATION SELECTION

ABSTRACT A computational aid for the diagnosis of genetic diseases based on statistical information is developed. The system operates on the basis of improving a diagnostic hypothesis by a suitable choice of the next physical examination or laboratory test for maximally reducing the hypothesis information entropy or amount of uncertainty. The method was applied to a reduced group of trisomies to test its behaviour on real patients. The main characteristics of the computational realization on a personal computer are described which constitute a friendly interface with the medical user.