

ORIGINALES

Concordancia diagnóstica entre atención primaria y atención especializada tras consulta urgente

G. Rodríguez Caravaca^a e I. Villar del Campo^b

Fundación Hospital Alcorcón y Centro de Salud Los Cármenes. Área 7. Madrid.

Objetivo. Evaluar la concordancia diagnóstica entre los medios intra y extrahospitalario tras consulta urgente.

Diseño. Estudio descriptivo transversal.

Emplazamiento. Área Sanitaria 7. Madrid.

Pacientes. Cualquier paciente que hubiera consultado de forma urgente en atención primaria y hubiera sido remitido a urgencias del hospital de referencia.

Mediciones y resultados principales. Se hizo una estimación muestral teniendo en cuenta una proporción de acuerdo esperado del 50%, una confianza del 95%, una precisión del 5% y una ausencia de diagnóstico en los partes de interconsulta del 30%. Se estimaron necesarios 551 casos. Se revisaron las historias clínicas del servicio de urgencias del hospital. Se codificaron los diagnósticos intra y extrahospitalarios según la clasificación internacional de enfermedades (CIE-9 MC). Se halló el índice kappa de concordancia entre diagnósticos individuales y el índice kappa global. Se revisaron los diagnósticos intra y extrahospitalarios de 559 pacientes. En 112 (20%) no se especificaba el diagnóstico extrahospitalario. El índice kappa fue muy alto ($\kappa \geq 0,8$) para la angina de pecho y urticaria. El índice kappa fue alto ($0,6 \geq \kappa < 0,8$) para ictus y trombosis venosa profunda (TVP); moderado ($0,4 \geq \kappa < 0,6$) para neumonía e insuficiencia cardíaca; bajo ($0,2 \geq \kappa < 0,4$) para apendicitis y artritis, y muy bajo ($\kappa < 0,2$) para meningitis y celulitis. El coeficiente kappa global para todos los diagnósticos fue de 0,65 (IC del 95%, 0,58-0,72).

Conclusiones. La concordancia global fue alta. La concordancia más elevada se obtuvo para procesos con diagnóstico fundamentalmente clínico, y la más baja para enfermedades que requieren exploraciones complementarias especializadas.

Palabras clave: Atención primaria; Concordancia; Dermatología; Hospital; Urgencias.

DIAGNOSTIC CONCORDANCE BETWEEN PRIMARY AND SPECIALIST CARE AFTER URGENT CONSULTATIONS

Objective. To evaluate the diagnostic concordance between primary health-care level and hospital health-care level after emergency visits.

Design. Cross-sectional study.

Setting. Health-care area 7 in Madrid.

Participants. Any patient studied in the primary health-care level and further sent by referral request to the emergencies of the health-care area reference hospital.

Measurements and results. The sample size was estimated according to a confidence level of 95%, a precision level of 5%, a concordance level of 50% and a 30% of referral requests without diagnosis. All diagnosis were codified by the ICD-9 CM. Single kappa index for each diagnosis and global kappa index were calculated.

559 patients were studied. 447 (80%) of the patients were referred with diagnosis and 112 (20%) without it. Kappa index was very high ($\geq 0,8$) for the following diagnosis: angina pectoris, and urticaria. Kappa index was high ($0,6 \geq \kappa < 0,8$) for stroke, and deep venous thrombosis. It was moderate ($0,4 \geq \kappa < 0,6$) for pneumonia, heart failure and heart attack. Kappa index was low ($0,2 \geq \kappa < 0,4$) for appendicitis, and arthritis, and it was very low ($< 0,2$) for meningitis, and cellulitis. Global kappa index was 0,65 (95% CI, 0,58-0,72).

Conclusions. Global concordance was high. The highest concordance was obtained for diseases with clinical diagnosis. Most of diseases with low and very low concordance are diseases that need specialized clinical tests.

(Aten Primaria 2000; 25: 292-296)

^aUnidad de Investigación. ^bMédico de Familia.

Correspondencia: G. Rodríguez Caravaca.

Unidad de Medicina Preventiva. Fundación Hospital Alcorcón.
Avda. Budapest, 1. 28925 Alcorcón (Madrid).

Manuscrito aceptado para su publicación el 6-IX-1999.

Introducción

La creación de la especialidad de medicina familiar y comunitaria¹, la formación de médicos de familia según un programa estructurado docente², el aumento de la cobertura poblacional por parte de equipos de atención primaria (EAP)³, el uso de pruebas diagnósticas en atención primaria (AP)⁴ y la aplicación de protocolos y programas de salud⁵ han supuesto un gran avance de la AP en nuestro país.

Este desarrollo de la AP ha permitido acercar la atención sanitaria al paciente y ofrecerle una de las características básicas de la misma: la continuidad de la atención, proporcionando asistencia médica de urgencia en el centro de salud, en el domicilio o en cualquier medio cuando ésta es solicitada por el paciente del sistema nacional de salud⁶. La resolución de los problemas urgentes en atención primaria de salud ha sido objeto de muchos estudios y trabajos y es un debate que continúa abierto⁷⁻⁹.

No obstante, hay que precisar que los medios de los que dispone la AP para resolver problemas urgentes son limitados, por lo que muchos de los pacientes que los aquejan acaban siendo derivados al hospital más cercano^{10,11}.

Esta situación nos permite estudiar la posible discordancia que puede surgir entre 2 observadores al evaluar un mismo fenómeno y que tiene que ver con la precisión o repetibilidad de las pruebas, en este caso, un paciente que solicita ayuda urgente y que es valorado en los medios extra e intrahospitalarios y al que se le diagnostica una enfermedad determinada.

El objetivo de este estudio es evaluar la concordancia diagnóstica entre los medios extrahospitalario (AP) y especializado (hospital de referencia) en pacientes que consultaron de for-

ma urgente en el Área Sanitaria 7 de la Comunidad de Madrid.

Material y métodos

Se diseñó un estudio descriptivo transversal que se llevó a cabo durante los meses de febrero a mayo de 1997 en el Servicio de Urgencias del Hospital Clínico San Carlos de Madrid, hospital de referencia del Área Sanitaria 7 de la Comunidad de Madrid. Se incluyó a cualquier paciente atendido en el servicio de urgencias del hospital que hubiera sido remitido al mismo mediante un parte de interconsulta tras demanda de asistencia urgente en su centro de salud y hubiese sido visto por su médico titular o por el médico de guardia del centro, en horario diurno y durante los días laborables. Se excluyó del estudio a cualquier enfermo que hubiera acudido directamente al hospital sin previa consulta con el nivel primario, al que se le hubiera remitido al hospital sin el parte de interconsulta o al que hubiera consultado por la noche, en día festivo o el fin de semana. Se hizo una estimación muestral teniendo en cuenta la situación más desfavorable de una proporción de acuerdo esperado del 50%, una confianza del 95%, una precisión del 5% y una falta de especificación diagnóstica en las partes de interconsulta del 30%, para lo que se estimaron necesarios 551 pacientes.

Se recogieron de cada paciente las variables sociodemográficas edad y sexo y el diagnóstico emitido en el servicio de urgencias del hospital y el reflejado en el parte de derivación del médico del nivel primario. Los diagnósticos eran posteriormente codificados por un único codificador, según la clasificación internacional de enfermedades CIE-9 MC, de forma independiente los diagnósticos de AP de los diagnósticos de urgencias. La codificación se realizó asignando los códigos según la máxima disagregación posible, según lo especificado en los partes de interconsulta y en el diagnóstico del informe de urgencias. La codificación de partes topográficas no se tuvo en cuenta, salvo que pudiera haber confusión o fuera una característica inherente al diagnóstico. Finalmente se realizó una recodificación de conveniencia, asignando a cada código CIE-9 MC un número consecutivo desde el 1 hasta el número total de diagnósticos estudiados, para reducir el número de dígitos de los diagnósticos codificados con la CIE-9 MC y facilitar así el manejo de los datos en la base de datos creada al efecto y en el programa estadístico-matemático.

Se describió la distribución de frecuencias de las variables cualitativas con sus porcentajes y sus intervalos de confianza (IC) del 95%. Se describieron las variables cuantitativas con la media y la desviación estándar, junto con sus IC del 95%. La asociación o independencia de las variables cualitativas se realizó con la prueba de la ji-cuadrado y la de las variables cuantitativas con la t de Student.

TABLA 1. Diagnósticos más frecuentes emitidos en el nivel de atención primaria y prevalencia en nuestra muestra

| Diagnóstico | N.º de casos | Prevalencia (%) |
|-------------------------------|--------------|-----------------|
| Neumonía | 28 | 5 |
| Apendicitis | 22 | 3,9 |
| Ictus | 21 | 3,8 |
| TVP* | 16 | 2,9 |
| Artritis aguda | 14 | 2,5 |
| Gastroenteritis aguda | 13 | 2,3 |
| Angina de pecho | 12 | 2,1 |
| Crisis hipertensiva | 11 | 2 |
| Broncospasmo | 10 | 1,8 |
| Infección urinaria | 10 | 1,8 |
| Urticaria | 9 | 1,6 |
| Dolor abdominal no quirúrgico | 9 | 1,6 |
| Sin diagnóstico | 112 | 20 |
| Resto diagnósticos | 272 | 48,6 |
| Total | 559 | 100 |

*TVP: trombosis venosa profunda.

Se calculó el índice kappa de concordancia individual, para cada diagnóstico emitido, con su IC del 95% y su significación estadística. También se halló el índice kappa global con su IC del 95% y su significación estadística.

Se diseñó una hoja de recogida de datos con el procesador de textos Word 7.0 para Windows. Para la gestión de los datos se elaboró una base de datos con la hoja de cálculo Excel 7.0 para Windows. El análisis estadístico y el tratamiento matemático se realizaron con el programa estadístico SAS.

Resultados

Se incluyó en el estudio un total de 559 pacientes. De ellos, 241 (43,2%) eran varones (IC del 95%, 38,1-47,8) y 318 (56,8%) mujeres (IC del 95%, 52,2-61,9), habiéndose encontrado diferencias estadísticamente significativas entre los porcentajes ($p < 0,05$). La edad media de los pacientes estudiados fue de 53,4 años (DE, 20), 55,6 años para los varones (IC del 95%, 53,5-57,8) y de 51,6 para las mujeres (IC del 95%, 49,7-53,5), habiéndose encontrado diferencias estadísticamente significativas entre la edad media según los sexos ($p < 0,05$). El rango de edad de los pacientes estudiados osciló en 15-92 años.

En 112 (20%) de los pacientes estudiados no se emitió ningún diagnóstico en el medio extrahospitalario. El número total de diagnósticos diferentes emitidos entre los dos niveles

asistenciales fue de 125. En AP se diagnosticó un total de 101 procesos diferentes y en especializada 104. Los diagnósticos más frecuentes emitidos por los facultativos del nivel primario fueron: neumonía, apendicitis, ictus, TVP y artritis aguda. La **tabla 1** muestra la frecuencia absoluta de los diagnósticos extrahospitalarios junto con su prevalencia en nuestra muestra. Los diagnósticos más frecuentes realizados por los médicos del nivel especializado fueron: patología crónica sin reagudización, ictus, estreñimiento, gastroenteritis aguda e infección del tracto urinario. La **tabla 2** muestra la frecuencia absoluta de los diagnósticos intrahospitalarios junto con su prevalencia en nuestra muestra. En la **figura 1** se puede apreciar la frecuencia absoluta de los diagnósticos más prevalentes en los 2 niveles asistenciales comparada entre sí.

Los principales diagnósticos, agrupados según los intervalos de clase habituales del índice kappa, se muestran en la **tabla 3**. Se consideró concordancia muy alta si $k \geq 0,8$, alta si $0,6 \geq k < 0,8$, moderada si $0,4 \geq k < 0,6$, baja si $0,2 \geq k < 0,4$ y muy baja si $k < 0,2$. La concordancia diagnóstica entre los dos niveles asistenciales para los diagnósticos más prevalentes con sus respectivos IC del 95% y su significación estadística se resumen en la **tabla 4**.

TABLA 2. Diagnósticos más frecuentes emitidos en el nivel especializado y prevalencia en nuestra muestra

| Diagnóstico | N.º de casos | Prevalencia (%) |
|---------------------------------|--------------|-----------------|
| Proceso crónico sin agudización | 24 | 4,3 |
| Ictus | 14 | 2,5 |
| Estreñimiento | 14 | 2,5 |
| Gastroenteritis aguda | 13 | 2,3 |
| Infección urinaria | 13 | 2,3 |
| Broncospasmo | 12 | 2,1 |
| Insuficiencia cardíaca | 12 | 2,1 |
| Neumonía | 12 | 2,1 |
| Dolor osteomuscular | 11 | 2,0 |
| Angina de pecho | 10 | 1,8 |
| Dolor abdominal no quirúrgico | 10 | 1,8 |
| TVP* | 10 | 1,8 |
| Resto diagnósticos | 404 | 72,4 |
| Total | 559 | 100 |

*TVP: trombosis venosa profunda.

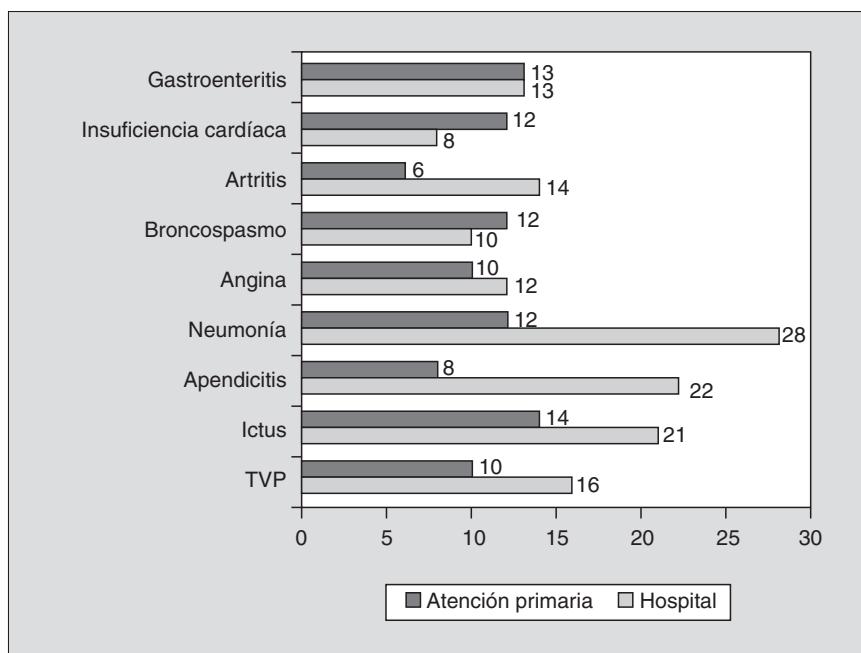


Figura 1. Comparación entre el número de diagnósticos más frecuentes emitidos según los dos niveles asistenciales.

La concordancia medida con el índice de kappa global fue alta, con kappa = 0,65 (IC del 95%, 0,58-0,72; $p < 0,001$).

Discusión

La fiabilidad (reproducibilidad, precisión, estabilidad o consistencia) de una medida hace referencia al grado en que pueden reproducirse los re-

sultados obtenidos por un procedimiento o instrumento de medición, en nuestro caso el diagnóstico clínico. Cuanta menor variabilidad se obtenga al repetir una medición, más fiable será el proceso. La fiabilidad no es una propiedad intrínseca de una medida, sino más bien de un instrumento cuando se aplica en una muestra de individuos en unas condiciones dadas¹².

La medida de la fiabilidad del diagnóstico clínico emitido por 2 o más observadores, en tanto que variable cualitativa, se puede medir con el índice kappa de Cohen, índice de medida adecuado porque elimina la fracción de concordancia debida la azar¹³. En nuestro estudio hemos considerado 2 observadores: por una parte el nivel de AP y por otra el nivel especializado de la urgencia del Hospital Clínico San Carlos, hospital de referencia del Área 7, área sanitaria de carácter urbano y considerando ambos observadores de forma global.

La variabilidad de diagnósticos emitidos por los médicos en nuestro estudio ha sido parecida y el número de diagnósticos realizados en los 2 niveles asistenciales ha sido prácticamente el mismo. El grado de concordancia global ha sido alto ($k = 0,65$; IC del 95%, 0,58-0,72), mayor que el de estudios similares realizados en el ámbito de otras especialidades médicas donde se compara la concordancia con el nivel de asistencia primaria^{14,15}, e incluso que el realizado con pruebas diagnósticas como el electrocardiograma comparando a médicos especialistas hospitalarios con médicos de atención primaria^{16,17}.

Habría que tener en cuenta además que el índice de concordancia kappa disminuye cuando la prevalencia es baja¹⁸, por lo que con un mayor número de casos de los diagnósticos estudiados nuestros resultados probablemente hubieran sido mejores.

Dentro de las categorías diagnósticas más prevalentes, los médicos de AP y del hospital diagnosticaron con suficiente precisión ictus, cólicos nefríticos, cólicos biliares e infecciones urinarias. Es de destacar la baja fiabilidad de algunos diagnósticos muy prevalentes, como apendicitis ($k = 0,36$) o la artrosis (tabla 3). Muchas enfermedades de las incluidas en el estudio no se han podido valorar adecuadamente y ofrecieron una concordancia muy baja debido al escaso número de pacientes estudiados.

No hay estudios publicados sobre la concordancia diagnóstica entre niveles asistenciales tras consulta urgente. Las diferentes condiciones de trabajo en las que se desarrolla la labor asistencial en los 2 ámbitos (presión asistencial, medios técnicos disponibles, tiempo que el profesional puede disponer para cada paciente, organización...) son factores

TABLA 3. Diagnósticos más frecuentes agrupados según el grado de concordancia diagnóstica para cada intervalo del índice kappa

| Intervalo del índice kappa | Diagnósticos más frecuentes en cada categoría |
|---|--|
| Concordancia muy alta, $k \geq 0,8$ | Angina de pecho, broncospasmo, crisis hipertensiva, gastroenteritis aguda, urticaria, crisis epiléptica |
| Concordancia alta, $0,6 \geq k < 0,8$ | Ictus, infección urinaria, TVP*, migraña, cólico nefrítico, cólico biliar |
| Concordancia moderada, $0,4 \geq k < 0,6$ | Dolor abdominal no quirúrgico, neumonía, estreñimiento, insuficiencia cardíaca, infarto agudo de miocardio |
| Concordancia baja, $0,2 \geq k < 0,4$ | Apendicitis, artritis aguda, dolor osteomuscular, artrosis |
| Concordancia muy baja, $k < 0,2$ | Meningitis, celulitis, tromboflebitis, crisis de ansiedad |

*TVP: trombosis venosa profunda.

TABLA 4. Índice kappa para los diagnósticos más prevalentes con su intervalo de confianza del 95% y su significación estadística

| Diagnóstico | Kappa | IC del 95% | p |
|-------------------------------|-------|------------|--------|
| Neumonía | 0,59 | 0,41-0,77 | < 0,05 |
| Apendicitis | 0,39 | 0,17-0,61 | < 0,05 |
| Ictus | 0,79 | 0,63-0,95 | < 0,05 |
| TVP* | 0,76 | 0,60-0,94 | < 0,05 |
| Infección urinaria | 0,69 | 0,47-0,91 | < 0,05 |
| Artritis aguda | 0,39 | 0,12-0,65 | < 0,05 |
| Gastroenteritis aguda | 0,84 | 0,70-0,98 | < 0,05 |
| Angina de pecho | 0,91 | 0,80-1,00 | < 0,05 |
| Dolor osteomuscular | 0,30 | 0,01-0,59 | < 0,05 |
| Crisis hipertensiva | 0,84 | 0,65-1,00 | < 0,05 |
| Dolor abdominal no quirúrgico | 0,41 | 0,14-0,68 | < 0,05 |
| estreñimiento | 0,59 | 0,34-0,84 | < 0,05 |
| Broncospasmo | 0,91 | 0,82-1,00 | < 0,05 |
| Insuficiencia cardíaca | 0,49 | 0,24-0,76 | < 0,05 |

*TVP: trombosis venosa profunda.

que podrían explicar las diferencias detectadas.

El estudio se ha desarrollado durante el invierno/primavera de 1997, por lo que desconocemos el efecto que la variación estacional haya podido tener en los resultados y puede ser una limitación del trabajo. Tampoco sabemos si el hecho de estudiar pacientes que han consultado sólo en horario diurno y en días laborales afecta en algún sentido nuestros resultados. Teniendo en cuenta que el Área 7 de Salud de Madrid es un área sanitaria de carácter urbano, y que los centros de salud no prestan asistencia continuada durante la noche, días festivos y fines de semana, sino que ésta es prestada por los servicios del 061, creemos que esta estrategia nos permite acercarnos a un conocimien-

to más adecuado de la concordancia diagnóstica entre el nivel especializado de urgencias y el nivel primario de los centros de salud de forma global, en sintonía con nuestro objetivo principal. Se ha intentado minimizar al máximo, dado que una sola persona codificó todos los diagnósticos, el sesgo de revisión diagnóstica, codificando de forma independiente los diagnósticos de AP y los diagnósticos de urgencias.

Se han encontrado diferencias estadísticamente significativas en la edad media entre los sexos y entre los porcentajes de varones y mujeres incluidos en el estudio. Esto podría indicar que la prevalencia de la patología urgente es diferente entre sexos, que la edad a la que se presenta varía o que los patrones que deter-

minan la consulta urgente son diferentes. Las características descriptivas de nuestro estudio no permiten concluir estas afirmaciones, lo que habría que evaluar con estudios experimentales.

Para profundizar en la repercusión de las diferentes condiciones de trabajo en los 2 niveles asistenciales, habría que diseñar estudios experimentales que tuvieran en cuenta las posibles variables de influencia. No obstante, como paso previo, pensamos que los resultados aquí obtenidos podrían ser usados por los médicos del nivel hospitalario y del nivel primario para plantear actividades conjuntas de formación continuada, elaboración de guías de actuación, sesiones, aumento de recursos, etc., tendentes a mejorar la homogeneidad de las actividades de su práctica clínica diaria.

Asimismo, pensamos que nuestros resultados pueden servir también para comparar los encontrados en nuestra área sanitaria con otras áreas de semejante estructura y organización.

Bibliografía

- Real Decreto 2.015/78 de 15 de julio y Real Decreto 3.303/78 de 29 de diciembre.
- Comisión Nacional de la especialidad. Programa de la especialidad de medicina familiar y comunitaria. Madrid: Ministerio de Sanidad y Consumo. Ministerio de Educación y Ciencia, 1994.
- Real Decreto 137/84 sobre estructuras básicas de salud.
- OMS. Serie de informes técnicos n.º 689. Ginebra: OMS, 1983.
- Abánades JC, Prieto A, Casado V, García L. Programas de salud. Documentación de apoyo para la planificación de programas. Valladolid: Servicio de Atención Primaria. Dirección General de Salud Pública y Atención Primaria. Consejería de Bienestar Social. Junta de Castilla y León, 1990.
- Martín A, Cano JF. Atención primaria de salud (3.ª ed.). Barcelona: Doyma, 1994.
- Llorente S, Alonso M, Buznego B. Papel de la atención primaria en la frecuenciación al servicio de urgencias de un hospital comarcal. Aten Primaria 1996; 18 (5): 243-247.
- Diego F, Franch J, Álvarez JC, Álvarez F, De Pablo ML, Villamar J. Urgencias hospitalarias en el área sanitaria de León. Estudio de la repercusión de la Atención Primaria. Aten Primaria 1990; 7 (1): 37-43.

9. Camp J. Papel de los centros de atención primaria en la descongestión de los hospitales. *Med Clin (Barc)* 1988; 91: 779-782.
10. Valls J, Segura L, Viñas M, Avendaño E. Urgencias en atención primaria y derivación de pacientes al hospital. *Aten Primaria* 1990; 7 (8): 593-594.
11. Botia F. Origen de las urgencias hospitalarias. Estudio retrospectivo. *Todo Hospital* 1989; 61: 31-35.
12. Domènec JM, Granero R. Diseño de investigaciones: pruebas de concordancia. Barcelona: Signo, 1997.
13. Sackett D, Haynes R, Tugwell P. Epidemiología clínica (1.^a ed.). Madrid: Díaz de Santos, 1989; 30-59.
14. Bolumar F, Ruiz MT, Hernández I, Pascual E. Reliability of the diagnosis of rheumatic conditions at the primary health care level. *J Rheumatol* 1994; 21: 2344-2348.
15. Sánchez M, Tovar J, Medina MA. Concordancia diagnóstica entre los médicos de atención primaria y los reumatólogos. *Aten Primaria* 1994; 13: 446-448.
16. García L, Alonso MT, Olmos O, Barrios M, Arias P, Carratalá FJ. Utilidad del electrocardiograma anual en el control de la hipertensión arterial. *Aten Primaria* 1990; 7: 551-556.
17. Moreno A, Iglesias P, Arbesu E. Grado de concordancia en la interpretación de electrocardiogramas. *Aten Primaria* 1995; 16: 187-191.
18. Grove WM, Andreasen NC, McDonald-Scott P, Grore WM, Andreasen NC, McDonald-Scott P et al. Reliability studies on psychiatric diagnosis. *Arch Gen Psychiatry* 1981; 38: 408-413.