

Reproducibilidad de un cuestionario telefónico sobre factores de riesgo asociados al comportamiento y las prácticas preventivas

Iñaki Galán^a / Fernando Rodríguez-Artalejo^b / Belén Zorilla^a

^aServicio de Epidemiología. Instituto de Salud Pública. Consejería de Sanidad de la Comunidad de Madrid. Madrid.

^bDepartamento de Medicina Preventiva y Salud Pública. Universidad Autónoma de Madrid. Madrid. España.

Correspondencia: Iñaki Galán. Servicio de Epidemiología. Instituto de Salud Pública. Consejería de Sanidad de la Comunidad de Madrid. Julián Camarillo, 4. 28037 Madrid. España.
Correo electrónico: inaki.galan@madrid.org

Recibido: 3 de julio de 2003.
Aceptado: 15 de diciembre de 2003.

(Reproducibility of a telephone questionnaire on risk factors associated with behavior and preventive practices)

Resumen

Objetivo: Se evalúa la reproducibilidad del cuestionario utilizado en un sistema de vigilancia de factores de riesgo de enfermedades no transmisibles (SIVFRENT), basado en encuestas continuas con entrevista telefónica.

Métodos: El mismo cuestionario fue aplicado telefónicamente en 2 ocasiones a una muestra de 586 personas de 18-64 años representativas de la Comunidad de Madrid, con un intervalo de 13-32 días. Se analizan 27 variables relacionadas con la percepción del estado de salud, antropometría, actividad física, consumo de tabaco y alcohol, prácticas preventivas, accidentalidad y seguridad vial. Para valorar la estabilidad de las respuestas se analizó la concordancia absoluta, los valores del índice kappa sin ponderar y ponderado, y el coeficiente de correlación intraclass (CCI), en las variables categóricas nominales, ordinales y cuantitativas, respectivamente.

Resultados: La tasa de respuesta a la segunda entrevista fue del 78,7%, y participaron 461 personas. La comparación de las prevalencias y promedios en las 2 entrevistas refleja que en 19 de los 27 indicadores analizados la desviación relativa fue inferior al 5%, y sólo 2 superaron el 20%. La concordancia absoluta excedió del 85% en 13 de las 20 variables categóricas, y los coeficientes kappa y CCI superaron el valor de 0,8 en 10 variables, en 13 oscilaron entre 0,6 y 0,8, en 3 entre 0,4 y 0,6, y sólo una fue inferior a 0,4. No se observó un patrón uniforme de variables sociodemográficas asociadas a la discordancia.

Conclusiones: Los resultados sugieren que el cuestionario telefónico utilizado en el SIVFRENT sobre factores de riesgo asociados al comportamiento y las prácticas preventivas es reproducible en una amplia muestra representativa de la población adulta de la Comunidad de Madrid.

Palabras clave: Reproducibilidad. Encuesta telefónica. Factores de riesgo.

Abstract

Objective: To assess the reproducibility of a questionnaire used in a surveillance system for risk factors of non-communicable diseases (SIVFRENT), based on continuous surveys through telephone interviews.

Methods: The same questionnaire was administered telephonically to a cross-section of 586 individuals aged between 18 and 64 years representative of the Autonomous Community of Madrid (Spain) on 2 occasions with an interval of 13 to 32 days. Twenty-seven variables related to perceived health status, anthropometry, physical activity, smoking, alcohol consumption, preventive practices, injuries and road safety were analyzed. To assess response stability, absolute agreement, unweighted and weighted kappa values, and the intraclass correlation coefficient (ICC) were used in nominal, ordinal and quantitative categorical variables, respectively.

Results: The response rate to the second interview involving 461 people was 78.7%. The comparison of prevalences and averages from both interviews revealed that in 19 out of the 27 indicators analyzed, the relative deviation was under 5% and exceeded 20% in only 2 indicators. Absolute agreement was greater than 85% in 13 of the 20 categorical variables. In 10 variables, the kappa and ICC coefficients were greater than 0.8, in 13 they ranged from 0.6 to 0.8, in 3 from 0.4 to 0.6 and only 1 was below 0.4. A uniform pattern of sociodemographic variables associated with disagreement was not observed.

Conclusions: The results suggest that the telephone questionnaire used in the surveillance system on risk factors associated with behavior and preventive practices (SIVFRENT) is reproducible in a wide sample representative of the adult population of the Autonomous Community of Madrid.

Key words: Reliability. Telephone survey. Risk factors.

Introducción

Las enfermedades crónicas, como las del corazón, las cerebrovasculares, el cáncer, la diabetes mellitus, la enfermedad pulmonar obstructiva crónica (EPOC), las del hígado y los accidentes, son la causa de casi el 70% de la mortalidad en España¹.

Actualmente, sabemos que gran parte de estos problemas de salud están relacionados con factores de riesgo modificables, como el consumo de tabaco, el consumo inadecuado de alcohol, el sedentarismo, la dieta desequilibrada, la obesidad o la hipertensión². En las últimas décadas, parte de la tendencia descendente observada en la incidencia y la mortalidad por cardiopatía isquémica y por enfermedades cerebrovasculares^{3,4}, así como por algunos tipos de cáncer⁵, parece estar relacionada con el mayor control de estos determinantes. En este contexto, el plan de salud de nuestra región recoge entre sus objetivos prioritarios la prevención y el control de estos factores⁶.

El Sistema de Vigilancia de Factores de Riesgo Asociados a Enfermedades no Transmisibles (SIVFRENT) se puso en marcha en 1995 en la Comunidad de Madrid para conocer la distribución y la evolución de los factores de riesgo relacionados con el comportamiento y las prácticas preventivas⁷. Su información sirve para elaborar las políticas de intervención en salud pública y evaluar su impacto en la población. Se basa en un sistema de encuestas continuas a personas de 18-64 años desarrolladas con metodología telefónica, siguiendo el modelo del Behavioral Risk Factor Surveillance System, establecido en 1984 en Estados Unidos por los Centers for Disease Control and Prevention (CDC). Este sistema recoge información mensual sobre factores de riesgo del comportamiento, prácticas preventivas y acceso al sistema sanitario, y está relacionado principalmente con la vigilancia de las enfermedades crónicas y accidentalidad. Actualmente, es la encuesta telefónica con mayor número de entrevistas del mundo⁸.

La utilización de encuestas telefónicas se ha incrementado notablemente en las últimas décadas debido, sobre todo, a la buena relación coste-eficacia y a la posibilidad de disponer rápidamente de la información recogida⁹. Sin embargo, en nuestro ámbito geográfico tenemos un conocimiento escaso acerca de su calidad. El estudio de la reproducibilidad de los instrumentos de medida es una práctica habitual en la investigación en salud pública, es decir, la capacidad, en este caso del cuestionario, de obtener la misma respuesta cuando se aplica en las mismas condiciones y a los mismos entrevistados. La estabilidad de las respuestas es un criterio importante para valorar la calidad de los datos de sistemas de vigilancia como el SIV-

FRENT, basados en encuestas continuas, objetivo que nos proponemos desarrollar en este estudio.

Métodos

El SIVFRENT es un sistema de vigilancia basado en encuestas continuas sobre los hábitos de salud y las prácticas preventivas en la población de 18-64 años no institucionalizada de la Comunidad de Madrid. La selección muestral se realiza a partir del directorio de hogares con telefonía fija, que en nuestra región supone el 94,8%¹⁰. Hay 2 etapas de muestreo: en primer lugar, se selecciona aleatoriamente el número de teléfono por muestreo estratificado según el ámbito geográfico y, una vez se ha contactado con el hogar, se selecciona al individuo a partir de 6 estratos de sexo y edad, según un procedimiento generado aleatoriamente por ordenador. Se establecen cuotas de sexo y edad proporcionales a su distribución en la población para facilitar su representatividad. Si el número de teléfono no contesta en un primer intento, se remarca un mínimo de 20 veces en diferentes franjas horarias y días. Las entrevistas se realizan durante una semana cada mes. El método de entrevista utilizado es el Computer Assisted Telephone Interviewing (CATI), que emplea un sistema interactivo mediante ordenador¹¹. Este sistema reduce la variabilidad entre entrevistadores y evita los fallos de registro debidos a filtros entre preguntas¹².

El cuestionario consta de un núcleo central de preguntas que se mantiene fijo desde 1995, año de la primera encuesta, y otro flexible que se modifica a propuesta de la estructura de salud pública. Para su elaboración se tomaron como referencia 5 encuestas de salud realizadas en España y 4 de otros países⁷.

La muestra de este estudio estuvo compuesta por las entrevistas realizadas durante los meses de septiembre, octubre y noviembre de 2002. Dado que el contacto es anónimo, a las personas encuestadas en la primera entrevista se les pidió permiso para volver a contactar con ellas posteriormente, sin informarles de que el cuestionario iba a ser repetido. El equipo de entrevistadores y el cuestionario utilizado fueron los mismos en ambas entrevistas, y el tiempo transcurrido entre las dos aplicaciones osciló entre 13 y 32 días.

Las variables analizadas fueron 27: la percepción del estado de salud (muy bueno, bueno, regular, malo y muy malo); peso y talla declarados y cálculo a partir de éstos del índice de masa corporal (kg/m^2); actividad física laboral habitual (4 categorías); actividad física en tiempo libre, estimada mediante la frecuencia y la duración de las actividades realizadas en las últimas 2 semanas, y calculando el gasto metabólico equivalente (MET)¹³ expresado como MET/min/semana (continua), y realización de dietas en los últimos 6 meses (2 ca-

tegorías). A los participantes también se les preguntó por el consumo de tabaco: haber fumado más de 100 cigarrillos y, en el caso de respuesta positiva, si es fumador diario, ocasional o ex fumador; edad de inicio del consumo regular (continua), y consejo médico para dejar de fumar (2 categorías). En el apartado de consumo de alcohol se estimó el número de días con algún consumo durante el último mes (7 categorías); cantidad de alcohol puro ingerido durante la semana anterior, a partir del recuerdo del consumo de diversas bebidas alcohólicas (continua); consumo excesivo en una misma ocasión, definido como la ingesta de 80 ml de alcohol puro en los varones y más de 60 ml en las mujeres (2 categorías), y edad de inicio del consumo regular (continua). También se recogió información sobre las prácticas preventivas: medición de la presión arterial (PA) alguna vez (2 categorías); tiempo desde la última medición de la presión arterial (5 categorías), PA elevada (informada por el médico, la enfermera o ambos); medición alguna vez del colesterol (2 categorías); tiempo transcurrido desde la última medición del colesterol (5 categorías); colesterol elevado informado por el médico (2 categorías), y en las mujeres tiempo transcurrido desde la última realización de citologías (5 categorías) y mamografías (5 categorías). Finalmente, se registra la accidentabilidad en los últimos 12 meses (2 categorías) y la utilización del cinturón de seguridad por ciudad y carretera (4 categorías, respectivamente).

El análisis estadístico se basó, en primer lugar, en la comparación de los promedios de las estimaciones en la primera y segunda entrevistas, calculando las desviaciones relativas (cociente de los promedios en las 2 entrevistas menos uno, expresado en porcentaje). Posteriormente, se analizaron las concordancias absolutas, es decir, la proporción de acuerdo sobre el total de personas entrevistadas. También se estimó el coeficiente de correlación intraclase (CCI) basado en modelos de efectos aleatorios en variables cuantitativas¹⁴, y el test de la kappa en las cualitativas¹⁵, con los intervalos de confianza (IC) del 95%. El test de la kappa se estimó sin ponderar en todas las variables cualitativas y, además, fue ponderado utilizando los pesos basados en las desviaciones cuadráticas en el análisis de concordancia de variables ordinales¹⁶. La elección de estos pesos se ha basado en la recomendación por su equivalencia al CCI cuando las distribuciones en los marginales son idénticas¹⁷, y por ser el método más utilizado en la práctica¹⁸. La concordancia fue establecida sobre la pregunta original con todas las categorías de respuesta, a excepción de la clase social y el índice de masa corporal. La clase social fue estimada a partir de la clasificación nacional de ocupaciones de 1994 según la propuesta de la Sociedad Española de Epidemiología¹⁹, agrupando de mayor a menor la posición socioeconómica en clases I, II, III, IVa, IVb y V, y el índice de masa corporal se estimó a partir del peso y la

talla declarados. Los valores de kappa mayores de 0,8 representan un acuerdo casi perfecto; entre 0,6 y 0,8, notable; entre 0,4 y 0,6, moderado; entre 0,2 y 0,4, regular, y por debajo de 0,2, malo¹⁶. La misma clasificación se utilizó para el CCI.

Por otro lado, se analizó la asociación entre las variables sociodemográficas y la discordancia, mediante modelos de regresión logística dicotómica multivariante no condicional²⁰. Las variables dependientes se definieron como la falta de acuerdo, es decir, la ausencia de concordancia en las categorías de respuesta en las 2 entrevistas respecto a 8 variables del cuestionario: percepción del estado de salud, actividad física en el medio laboral, realización de dietas, haber fumado más de 100 cigarrillos en la vida, consumo de alcohol en el último mes, medición de la PA, medición del colesterol y accidentes en los últimos 12 meses. Estas variables fueron seleccionadas entre las 20 categóricas según 2 criterios: a) reducir la discordancia por cambios en el comportamiento entre las 2 entrevistas (p. ej., se excluye la realización de actividad física en tiempo libre en las últimas 2 semanas o el consumo de alcohol durante la última semana), y b) que estuvieran dirigidas a todas las unidades muestrales para no perder poder estadístico. Las variables independientes fueron la edad (3 categorías), el sexo, el nivel educativo (4 categorías) y la clase social (trabajadores manuales y no manuales), recogidas en la primera entrevista, controladas a su vez por el tiempo transcurrido entre las 2 entrevistas.

Resultados

La tasa de respuesta a la primera entrevista, definida como el cociente entre el número de entrevistas realizadas y el sumatorio de entrevistas realizadas, negativas de hogares, negativas de individuos y entrevistas incompletas fue del 65,3%. La distribución por edad y sexo es similar a la estructura de población de 18-64 años de la Comunidad de Madrid, aunque el nivel de estudios de la encuesta telefónica es moderadamente más alto. La tasa de respuesta a la segunda entrevista fue del 78,7%, y participaron en total 461 personas. Como se puede observar en la tabla 1, la distribución de los participantes en las 2 entrevistas es muy similar respecto a las variables sociodemográficas, y no existen diferencias estadísticamente significativas. Aun así, en la segunda entrevista se observa en los participantes una proporción ligeramente superior de mujeres, de mayor edad y con estudios inferiores, aunque estas diferencias no son estadísticamente significativas. Sin embargo, teniendo en cuenta la alta participación en la segunda entrevista y que las desviaciones comentadas anteriormente no son muy relevantes, la dis-

Tabla 1. Características sociodemográficas de los participantes en el estudio

	Total de participantes en la primera entrevista, n (%)	Participantes en la segunda entrevista, n (%)	No participantes en la segunda entrevista, n (%)
Sexo			
Varones	286 (48,8)	231 (50,1)	55 (44,0)
Mujeres	300 (51,2)	230 (49,9)	70 (56,0)
Edad (años)			
18-29	189 (32,3)	154 (33,4)	35 (28,0)
30-44	198 (33,8)	161 (34,9)	37 (29,6)
45-64	199 (34,0)	146 (31,7)	53 (42,4)
Estudios			
Primarios o inferiores	40 (6,8)	27 (5,9)	13 (10,4)
Secundarios de primer grado	141 (24,2)	109 (23,6)	33 (26,4)
Secundarios de segundo grado	212 (36,2)	172 (37,3)	40 (32,0)
Universitarios	192 (32,8)	153 (33,2)	39 (31,2)
Clase social*			
I	75 (14,0)	58 (13,6)	17 (15,6)
II	80 (15,0)	63 (14,8)	17 (15,6)
III	174 (32,6)	134 (31,5)	40 (36,7)
IVa	77 (14,4)	68 (16,0)	9 (8,3)
IVb	62 (11,6)	48 (11,3)	14 (12,8)
V	66 (12,4)	54 (12,7)	12 (11,0)

*Clase social I: mayor nivel; clase V: menor nivel.

tribución de los participantes en las 2 entrevistas es comparable a la de la muestra inicial.

La comparabilidad de las estimaciones de estos factores entre la primera y la segunda entrevistas es muy consistente: 19 de los 27 indicadores presentados en la tabla 2 tienen desviaciones relativas inferiores al 5%; 6, entre un 5 y un 20% (percepción del estado de salud, realización de dietas, consejo médico para dejar de fumar, consumo excesivo de alcohol en una misma ocasión, accidentes en los últimos 12 meses y no utilización del cinturón de seguridad por ciudad), y solamente 2 superan el 20%: la actividad física durante el tiempo libre (MET/min/semana) y conducir bajo los efectos del alcohol. No se observa ningún patrón general en la dirección de estas diferencias.

Los resultados de la concordancia individual se exponen en las tablas 3 y 4. La concordancia de las variables sociodemográficas es muy elevada. En el caso de la edad, es prácticamente perfecta y no se produce ninguna discordancia respecto al sexo. El nivel de estudios y la clase social, aunque ésta última en menor medida, reflejan una notable correlación tanto en los valores absolutos como en los coeficientes kappa (tabla 3).

La reproducibilidad de los factores de riesgo asociados al comportamiento y las prácticas preventivas también fue globalmente alta (tabla 4 y figura 1). La concordancia absoluta excede del 85% en 13 de las 20 variables categóricas. La concordancia más alta correspondió a la utilización del cinturón de seguridad por

ciudad (98,1%; IC del 95%, 96,2-99,2), y haber fumado más de 100 cigarrillos en la vida (97,2%; IC del 95%, 95,2-98,5). Por el contrario, las preguntas con menor acuerdo fueron el tiempo transcurrido desde la última medición del colesterol (62,2%; IC del 95%, 57,4-66,8) y desde la última medición de la PA (64,5%; IC del 95%, 59,9-69,0).

Después de corregir el acuerdo esperado por efecto del azar, los valores de kappa sin ponderar para las 20 variables categóricas fueron los siguientes: 2 variables tienen una concordancia casi perfecta (kappa superiores a 0,80); 8, notable (entre 0,60 y 0,80); 8, moderada (entre 0,40 y 0,60) (percepción del estado de salud, realización de dietas, consumo de alcohol en el último mes, medición del colesterol alguna vez, tiempo desde la última medición del colesterol, accidentes en los últimos 12 meses, no utilización de forma habitual del cinturón de seguridad por ciudad y por carretera), y 2, regular (entre 0,20 y 0,40) (medición de la PA alguna vez y tiempo desde la última medición). Los valores de kappa más altos sin ponderar correspondieron a las preguntas sobre consumo de tabaco y los más bajos, a la medición de la PA. De las variables ordinales, 3 tienen valores de kappa ponderados superiores a 0,80, y las 6 restantes oscilan entre 0,605 y 0,779. Finalmente, la concordancia observada en las variables cuantitativas es muy alta, con CCI superiores a 0,80 en 5 de las 7 variables registradas. Los mayores coeficientes corresponden a las variables antropométricas: 0,986 (IC del 95%, 0,985-0,988) para el peso y 0,975 (IC del 95%, 0,970-0,979)

Tabla 2. Estimaciones de las prevalencias y los promedios de factores relacionados con el comportamiento y las prácticas preventivas en la primera y segunda entrevistas

	Personas que participan en la primera y la segunda entrevistas		Respuesta de los que no participan en la segunda entrevista % o media (IC del 95%)
	Respuesta a la primera entrevista % o media (IC del 95%)	Respuesta a la segunda entrevista % o media (IC del 95%)	
Percepción del estado de salud			
Percepción de salud regular, mala o muy mala (%)	15,0 (11,8-18,6)	15,8 (12,6-19,5)	20,8 (14,1-29,0)
Antropometría			
Peso (media)	69,4 (68,2-70,6)	69,6 (68,3-70,8)	69,0 (66,8-71,1)
Estatura (media)	169,1 (168,2-169,9)	168,8 (168,0-169,7)	167,5 (165,8-169,2)
IMC (media)	24,2 (23,9-24,5)	24,3 (24,0-24,7)	24,5 (23,9-25,1)
Actividad física			
Sedentarios en la actividad habitual/laboral (%)	48,0 (43,4-52,6)	48,6 (43,9-53,3)	45,6 (36,7-54,7)
Actividad física en tiempo libre (MET/min/semana) (media)	1.243,6 (1.076,6-1.410,6)	1.030,9 (911,7-1.150,1)	828,2 (665,1-991,3)
Realización de dietas (%)	12,8 (9,9-16,2)	11,3 (8,5-14,5)	8,0 (3,9-14,2)
Consumo de tabaco			
Haber fumado > 100 cigarrillos (%)	55,7 (51,2-60,3)	56,4 (51,9-60,9)	58,4 (49,6-67,2)
Fumadores diarios (%)	32,1 (27,8-36,4)	32,5 (28,3-37,0)	36,8 (28,4-45,9)
Edad de inicio (media)	17,7 (17,0-18,5)	17,3 (16,7-17,9)	18,3 (16,9-19,6)
Consejo médico para dejar de fumar (%)	45,3 (37,1-53,7)	48,0 (39,8-56,3)	47,8 (32,9-63,1)
Consumo de alcohol			
Bebedores habituales (%)	62,0 (57,4-66,5)	62,3 (57,7-66,7)	62,4 (53,3-70,9)
Cantidad consumida en la semana anterior (%)	78,2 (66,9-89,5)	79,0 (67,5-90,5)	100,0 (70,0-130,0)
Consumo excesivo en una misma ocasión (%)	15,2 (12,0-18,8)	14,1 (11,1-17,6)	11,2 (6,3-18,1)
Edad de inicio (media)	18,3 (17,7-19,0)	17,9 (17,4-18,5)	19,2 (18,3-20,1)
Conducir bajo efectos del alcohol (%)	3,7 (2,2-5,8)	2,6 (1,4-4,5)	2,4 (0,5-6,9)
Prácticas preventivas			
Medición tensión arterial (%)	97,2 (95,2-98,5)	97,4 (95,5-98,6)	93,6 (87,8-97,2)
Medición de la presión arterial < 2 años (%)	88,9 (85,7-91,7)	87,2 (83,8-90,1)	85,6 (78,2-91,2)
Presión arterial elevada (%)	12,6 (9,7-16,0)	12,8 (9,9-16,2)	15,2 (9,4-22,7)
Medición del colesterol (%)	93,9 (91,3-95,9)	95,4 (93,1-97,2)	88,8 (81,9-93,7)
Medición del colesterol < 2 años (%)	84,2 (80,5-87,4)	83,9 (80,3-87,2)	77,6 (69,3-84,6)
Colesterol elevado (%)	15,0 (11,8-18,6)	14,8 (11,6-18,3)	24,8 (17,5-33,3)
Realización de citologías < 5 años (%)	74,3 (68,2-79,9)	74,3 (68,2-79,9)	77,1 (65,6-86,3)
Realización de mamografías < 5 años (%)	47,8 (41,2-54,5)	45,7 (39,1-52,3)	62,9 (50,5-74,1)
Accidentabilidad			
Accidentes en los últimos 12 meses (%)	8,0 (5,7-10,9)	8,9 (6,5-11,9)	5,6 (2,3-11,2)
Seguridad vial			
Uso no habitual cinturón por ciudad (%)	16,7 (13,3-20,5)	15,0 (11,7-18,3)	10,5 (5,6-17,7)
Uso no habitual del cinturón por carretera (%)	0,7 (0,1-2,1)	0,7 (0,1-2,0)	0,9 (0,02-5,1)

IC: intervalo de confianza; IMC: índice de masa corporal; MET: gasto metabólico equivalente.

Tabla 3. Concordancia de las variables sociodemográficas

	Concordancia absoluta % (IC del 95%)	Coefficientes de concordancia (IC del 95%)	Test
Edad (cont)		0,998 (0,998-0,999)	CCI
Sexo (cat 2)	100	1	κ
Estudios (cat 8)	84,4 (81,1-87,7)	0,789 (0,738-0,840)	κ
		0,941 (0,850-1)	κp
Clase social (cat 8)	68,6 (64,0-73,2)	0,628 (0,587-0,670)	κ
		0,796 (0,697-0,894)	κp

IC: intervalo de confianza; cont: variable continua; cat: variable categórica y número de categorías; κ: kappa sin ponderar; κp: kappa ponderado; CCI: coeficiente de correlación intraclase.

para la estatura, y el más bajo, a la estimación del gasto de energía en tiempo libre (0,649; IC del 95%, 0,591-0,700).

En la tabla 5 se describe la asociación entre las características sociodemográficas y la discordancia entre la primera y la segunda entrevistas en diversos factores estudiados. En general, no se observa un patrón uniforme en las 8 preguntas analizadas. En la percepción del estado de salud, las mujeres presentan una mayor discordancia, estadísticamente significativa, respecto a los varones, con una *odds ratio* (OR) de 1,56 (IC del 95%, 1,00-2,45; p = 0,048), y los individuos con estudios primarios o inferiores presentan una menor dis-

Tabla 4. Concordancia de las preguntas sobre factores relacionados con el comportamiento y las prácticas preventivas

	Concordancia absoluta % (IC del 95%)	Coefficientes de concordancia (IC del 95%)	Test
Percepción del estado de salud (cat ordinal 5)	74,2 (69,9-78,1)	0,477 (0,412-0,542) 0,605 (0,515-0,895)	κ κρ
Antropometría			
Peso (cont)		0,986 (0,985-0,988)	CCI
Estatura (cont)		0,975 (0,970-0,979)	CCI
IMC (cont)		0,964 (0,957-0,970)	CCI
Actividad física			
Actividad física laboral (cat ordinal 4)	81,9 (78,0-85,3)	0,697 (0,626-0,767) 0,779 (0,687-0,870)	κ κρ
Actividad física en tiempo libre (MET/min/semana) (cont)		0,649 (0,591-0,700)	CCI
Realización de dietas			
Realización de dietas en los últimos 6 meses (cat 2)	88,1 (84,8-90,9)	0,437 (0,346-0,528)	κ
Consumo de tabaco			
Haber fumado > 100 cigarrillos (cat 2)	97,2 (95,2-98,5)	0,943 (0,852-1)	κ
Consumo de tabaco (cat 4)	92,1 (88,0-95,1)	0,854 (0,754-0,953)	κ
Edad de inicio del consumo regular (cont)		0,824 (0,762-0,871)	CCI
Consejo médico para dejar de fumar (cat 2)	85,2 (78,3-90,6)	0,704 (0,539-0,868)	κ
Consumo de alcohol			
Consumo de alcohol último mes (cat ordinal 7)	65,1 (60,5-69,4)	0,572 (0,530-0,614) 0,840 (0,749-0,931)	κ κρ
Cantidad consumida en la semana anterior (cont)		0,863 (0,765-0,960)	CCI
Consumo excesivo en una misma ocasión en los últimos 30 días (cat 2)	87,8 (83,8-91,1)	0,616 (0,509-0,722)	κ
Edad de inicio del consumo regular (cont)		0,785 (0,739-0,824)	CCI
Haber conducido bajo los efectos del alcohol en los últimos 12 meses (cat 2)	88,7 (85,5-91,5)	0,692 (0,612-0,773)	κ
Prácticas preventivas			
Medición tensión arterial alguna vez (cat 2)	96,3 (94,2-97,8)	0,301 (0,210-0,392)	κ
Tiempo desde la última medición de la presión arterial (cat ordinal 5)	64,5 (59,9-69,0)	0,394 (0,345-0,453) 0,612 (0,519-0,705)	κ κρ
Presión arterial elevada (cat 4)	92,7 (89,9-95,0)	0,695 (0,620-0,770)	κ
Medición del colesterol alguna vez (cat 2)	95,4 (93,1-97,2)	0,548 (0,458-0,638)	κ
Tiempo desde la última medición del colesterol (cat ordinal 5)	62,2 (57,4-66,8)	0,438 (0,379-0,496) 0,672 (0,577-0,766)	κ κρ
Colesterol elevado (cat 2)	92,3 (89,3-94,6)	0,713 (0,618-0,808)	κ
Tiempo desde la última realización de citologías (cat ordinal 5)	83,0 (77,6-87,7)	0,753 (0,677-0,829) 0,886 (0,757-1)	κ κρ
Tiempo desde la última realización de mamografías (cat ordinal 5)	85,7 (80,4-89,9)	0,786 (0,705-0,866) 0,847 (0,718-0,977)	κ κρ
Accidentalidad			
Accidentes en los últimos 12 meses (cat 2)	92,6 (89,8-94,8)	0,524 (0,433-0,615)	κ
Seguridad vial			
Uso no habitual del cinturón por ciudad (cat ordinal 4)	80,0 (76,0-83,7)	0,548 (0,484-0,611) 0,745 (0,651-0,839)	κ κρ
Uso no habitual del cinturón por carretera (cat ordinal 4)	98,1 (96,2-99,2)	0,493 (0,416-0,569) 0,639 (0,545-0,733)	κ κρ

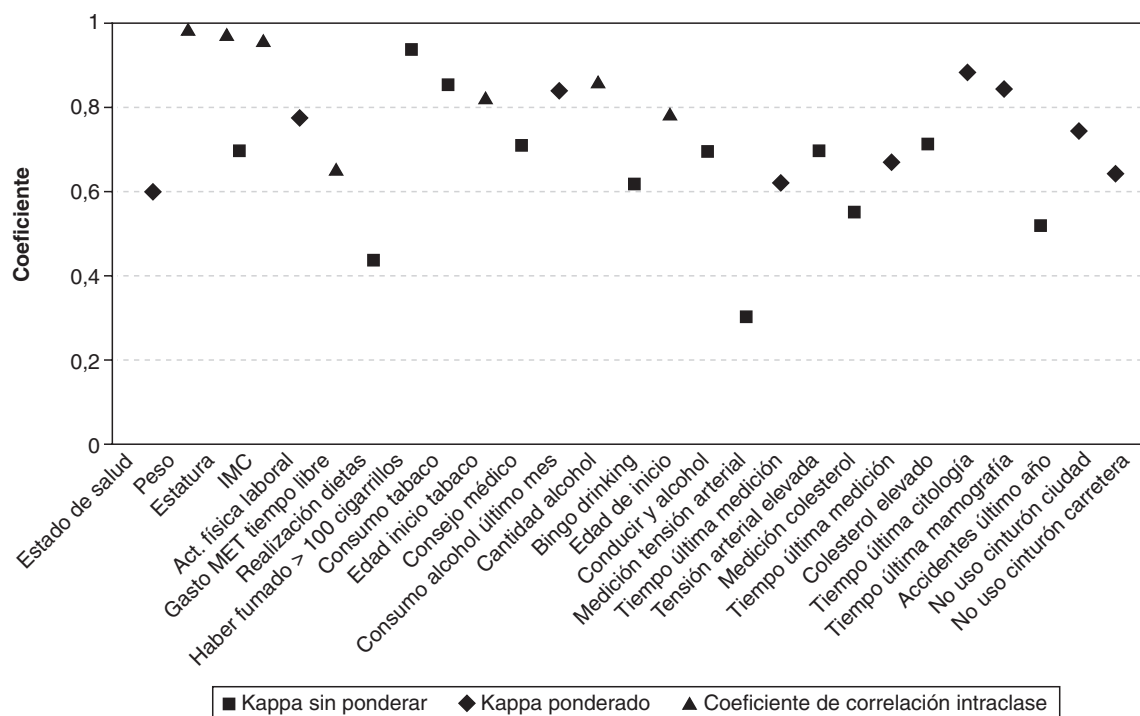
Cont: variable continua; cat: variable categórica y número de categorías.

κ: kappa sin ponderar; κρ: kappa ponderado; CCI: coeficiente de correlación intraclass; IMC: índice de masa corporal; MET: gasto metabólico equivalente.

cordancia que los que tienen estudios universitarios (OR = 0,18; IC del 95%, 0,04-0,87; p = 0,033). En la actividad física durante la actividad laboral, las personas de 45-64 años tienen una menor discordancia que las de 18-29 años (OR = 0,41; 0,19-0,86; p = 0,019), y también se observan discordancias estadísticamente sig-

nificativas en los que tienen menor nivel de estudios (OR = 9,68; IC del 95%, 3,07-30,47; p = 0,000). En la medición del colesterol, se describen menores discordancias en las mujeres respecto a los varones (OR = 0,18; IC del 95%, 0,05-0,66; p = 0,010), así como en las edades de 45-64 años frente a las personas más

Figura 1. Coeficientes de concordancia de 27 factores relacionados con el comportamiento y las prácticas preventivas.



jóvenes (OR = 0,15; IC del 95%, 0,03-0,80; p = 0,026). No se observan diferencias estadísticamente significativas en el resto de variables. Por otro lado, no se encuentran diferencias en la discordancia en función del tiempo transcurrido entre las 2 entrevistas.

Discusión

Los resultados de este estudio muestran una buena reproducibilidad del cuestionario sobre factores de riesgo asociados al comportamiento y las prácticas preventivas utilizado en el SIVFRENT. Estos resultados concuerdan con los observados en otros estudios, que utilizan también el método de encuestas continuas basadas en la metodología telefónica como fuente de información en los sistemas de vigilancia sobre estos factores²¹⁻²⁴.

Entre las variables sociodemográficas, el sexo y la edad presentan una concordancia perfecta o casi perfecta. El nivel de estudios y la clase social tienen una alta concordancia, sobre todo teniendo en cuenta el elevado número de categorías estudiadas. Estos resultados son coincidentes con otros de características metodológicas similares²¹⁻²⁵.

La percepción del estado de salud refleja una concordancia moderada (kappa sin ponderar = 0,477), la segunda concordancia absoluta más baja (74,2%). Este acuerdo es muy parecido al estimado por Martikainen et al en Finlandia, con valores de kappa próximos a 0,5²⁶, pero ligeramente más bajos que los observados por Lundberg et al en Suecia, aunque éstos reducen a 3 las categorías de respuesta²⁷. Esta variación entre las 2 entrevistas podría reflejar cambios reales en el propio estado de salud, ya que esta medida, a pesar de que se refiere a la valoración durante el último año, se puede ver afectada por los cambios causados por trastornos físicos y emocionales transitorios. La interpretación de la salud percibida varía según el sexo. En las mujeres, el peor estado de salud se asocia más a menudo con procesos dolorosos, mientras que en los varones se relaciona más con la limitación de la capacidad funcional²⁸. Esta interpretación podría justificar la mayor discordancia observada en mujeres, ya que los procesos dolorosos pueden tener mayor variabilidad que los de limitación física y ser producto de cambios reales en el tiempo. También se observa una mayor discordancia en las personas de mayor edad respecto a las más jóvenes, aunque no es de gran magnitud y las diferencias no son estadísticamente significativas. Este menor acuerdo podría relacionarse con la menor es-

Tabla 5. Características sociodemográficas asociadas a la discordancia entre la primera y la segunda entrevista

	Percepción del estado de salud OR (IC del 95%)	Actividad física laboral OR (IC del 95%)	Realización de dietas OR (IC del 95%)	Haber fumado >100 cigarrillos OR (IC del 95%)	Consumo de alcohol último mes OR (IC del 95%)	Medición de la presión arterial OR (IC del 95%)	Medición del colesterol OR (IC del 95%)	Accidentes durante los últimos 12 meses OR (IC del 95%)
Sexo								
Varones	1	1	1	1	1	1	1	1
Mujeres	1,56 (1,00-2,45) ^a	0,98 (0,58-1,66)	1,28 (0,70-2,32)	1,24 (0,40-3,83)	1,31 (0,87-1,98)	0,55 (0,18-1,71)	0,18 (0,05-0,66) ^a	0,57 (0,27-1,23)
Edad (años)								
18-29	1	1	1	1	1	1	1	1
30-44	1,16 (0,67-2,01)	0,62 (0,33-1,17)	0,50 (0,23-1,06)	0,49 (0,11-2,13)	1,26 (0,75-2,12)	0,60 (0,17-2,07)	0,87 (0,30-2,49)	0,84 (0,35-1,98)
45-64	1,24 (0,67-2,29)	0,41 (0,19-0,86) ^a	0,72 (0,33-1,58)	1,15 (0,29-4,56)	1,65 (0,93-2,92)	0,27 (0,05-1,45)	0,15 (0,03-0,80) ^a	0,43 (0,14-1,31)
Estudios								
Universitarios	1	1	1	1	1	1	1	1
Secundarios de segundo grado	0,67 (0,39-1,15)	1,20 (0,57-2,50)	1,35 (0,63-2,88)	0,94 (0,25-3,60)	0,73 (0,44-1,22)	1,26 (0,32-4,92)	0,80 (0,21-3,03)	1,83 (0,74-4,52)
Secundarios de primer grado	0,84 (0,43-1,62)	3,37 (1,49-7,60) ^a	2,20 (0,90-5,37)	0,71 (0,13-3,90)	0,77 (0,41-1,44)	1,31 (0,22-7,68)	2,22 (0,50-9,79)	0,81 (0,21-3,15)
Primarios o inferiores	0,18 (0,04-0,87) ^a	9,68 (3,07-30,47) ^b	1,73 (0,40-7,55)	- ^c	0,64 (0,23-1,78)	6,54 (0,75-57,25)	2,09 (0,16-27,83)	3,80 (0,73-19,84)
Clase social								
Trabajadores no manuales	1	1	1	1	1	1	1	1
Trabajadores manuales	1,10 (0,65-1,88)	1,74 (0,96-3,16)	0,58 (0,29-1,18)	1,11 (0,29-4,23)	1,05 (0,64-1,72)	1,03 (0,29-3,57)	1,55 (0,47-5,09)	0,78 (0,32-1,88)
Intervalo entre entrevistas								
13-19 días	1	1	1	1	1	1	1	1
20-24 días	0,86 (0,50-1,48)	1,06 (0,56-2,00)	0,86 (0,43-1,72)	1,80 (0,37-8,70)	1,34 (0,80-2,25)	1,01 (0,24-4,33)	0,58 (0,21-1,62)	1,89 (0,72-4,99)
25-32 días	1,07 (0,53-2,16)	1,31 (0,58-2,97)	0,59 (0,21-1,63)	1,52 (0,21-11,17)	1,44 (0,73-2,81)	3,92 (0,91-16,94)	0,41 (0,08-2,07)	0,81 (0,19-3,41)

OR: *odds ratio* estimadas por regresión logística, incluidas las variables sexo, edad, estudios, clase social e intervalo entre entrevistas.
^ap < 0,05; ^bp < 0,01; ^cno calculado debido a ausencia de datos en la categoría.

pecificidad de los problemas de salud crónicos en las personas de más edad²⁶. Por otro lado, se observa una menor discordancia en los niveles educativos más bajos, que son precisamente los que tienen peor estado de salud. Esto podría relacionarse con una mayor especificidad de la respuesta en esta población, ya que la mayor variabilidad en las categorías del estado de salud se produce entre las categorías de respuesta muy bueno y bueno. Los resultados de otros estudios no son del todo concordantes con nuestros datos: Martikainen et al²⁶ y Lundberg et al²⁷ no observan diferencias de concordancia por sexo, aunque sí detectan una menor correlación en la población de mayor edad, en ambos sexos²⁶ o sólo en mujeres²⁷. Por tanto, la explicación de las diferencias observadas en este aspecto no es concluyente.

El peso y la estatura declarados reflejan una concordancia muy alta, la mejor de todos los apartados, y ligeramente superior a la obtenida en estudios de referencia realizados con metodología similar al SIVFRENT^{21,23,24}.

El apartado de actividad física describe también una buena concordancia, más alta para la realizada en la actividad laboral que en el tiempo libre, posiblemente porque la primera se refiere a la frecuencia habitual y es relativamente estable en el tiempo, mientras que la segunda se basa en el recuerdo de la realización de actividades en las últimas 2 semanas y, por consiguiente, puede variar a causa de haberse producido verdaderos cambios entre las 2 entrevistas. De hecho, la actividad física en el tiempo libre tiene un fuerte componente estacional, con un incremento de su práctica durante los meses de verano²⁹. En nuestro estudio, al analizar el consumo de energía por meses, observamos que el gasto en septiembre es un 17% superior al de octubre y noviembre. Es más, si solamente analizamos la concordancia de las entrevistas correspondientes a octubre y noviembre, el CCI del gasto en MET se incrementa de 0,649 a 0,713 (IC del 95%, 0,647-0,768). En la cohorte EPIC, se observa también una moderada-buena correlación entre repeticiones

del cuestionario para las actividades de tiempo libre basado también en el cálculo de MET, aunque el período entre entrevistas es superior^{30,31}. Respecto a la actividad física laboral, Batty observa en los varones de mediana edad, con una pregunta similar a la utilizada por nosotros, un índice kappa de 0,56, ligeramente más bajo que en nuestro estudio³². En el análisis de la discordancia, observamos un menor acuerdo en los individuos que tienen menor nivel de estudios respecto a los universitarios en la actividad física laboral. Estos datos podrían estar relacionados con la menor estabilidad que tienen las mediciones en las actividades más vigorosas³¹, que son realizadas principalmente por los grupos poblacionales con menor nivel educativo y de clase social más baja. En este sentido, también los trabajadores manuales presentan mayor discordancia, aunque sus diferencias no llegan a ser estadísticamente significativas.

La realización de dietas tiene una concordancia absoluta superior al 80%, aunque una moderada concordancia al corregir por la concordancia esperada por azar, y no se observa una relación estadísticamente significativa entre la discordancia y las variables demográficas analizadas. Aunque no disponemos de referencias para comparar directamente estos resultados, Shea et al²⁴ observan un índice kappa más alto (0,65) en la realización de dietas para perder peso.

El consumo de tabaco es el apartado, después de la antropometría, que obtiene globalmente una mayor concordancia, al igual que se observa en diversos estudios de referencia, que describen concordancias elevadas^{23,24,33}.

El consumo de alcohol refiere concordancias más bajas que el de tabaco, aunque lo suficientemente estables teniendo en cuenta que se refieren a recuerdos de consumo a corto plazo. Cabe destacar que se obtiene una mayor concordancia en el cálculo del consumo de alcohol puro durante la última semana que la frecuencia de los días con consumo de alcohol en el último mes (agrupados en 7 categorías). Starr et al²², en un estudio de similares características realizado en Australia, estiman también unas concordancias muy parecidas, ya que obtienen un CCI de 0,86 para la frecuencia de consumo semanal de alcohol y un índice kappa ponderado de 0,75 para la clasificación de riesgo de alcohol agrupado en 4 categorías. Por otro lado, Stein et al²⁵ observan un índice kappa muy similar al estimado en nuestro estudio respecto al elevado consumo de alcohol, aunque obtienen concordancias más bajas en el indicador de conducir después de haber consumido alcohol de forma excesiva.

No se observa ninguna asociación entre las variables sociodemográficas y la discordancia respecto a la realización de dietas, consumo de tabaco y de alcohol. Stein et al²⁵ no detectan tampoco una relación respecto al consumo de tabaco, aunque sí una mayor discor-

dancia en el consumo de alcohol en la población más joven, en mujeres y personas de bajos ingresos económicos.

Respecto a las prácticas preventivas, la menor concordancia, al analizar conjuntamente la observada y los coeficientes kappa, corresponde al tiempo transcurrido desde la realización de la última medición de la PA y del colesterol. Sin embargo, es más alta la concordancia respecto al conocimiento de tener la PA o el colesterol elevados. Otros autores^{21,23-25} observan también menos estabilidad en la pregunta correspondiente a la realización de una revisión respecto al conocimiento que tienen sobre la PA o el colesterol elevados. Tanto en la medición de la PA como en el colesterol, las mujeres y las personas de mayor edad tienen una menor discordancia respecto a estas preguntas, aunque solamente se alcanza significación estadística en la medición del colesterol. Similares resultados observan Stein et al²⁵, aunque obtienen una menor proporción de discordancia, con significación estadística solamente en los grupos de mayor edad. Esta circunstancia podría estar relacionada con el mejor cumplimiento de las mujeres con los exámenes periódicos de salud, al igual que las personas de mayor edad respecto a las más jóvenes³⁴.

La concordancia del tiempo transcurrido desde la realización de mamografías y citologías es muy similar y considerablemente alta, teniendo en cuenta que constan de 5 categorías de respuesta. Browson et al²³ observan altas concordancias en esta pregunta, aunque definida en su estudio como las personas que se han realizado estas pruebas alguna vez y en el último año. También Stein et al³⁵ observan altas concordancias respecto a la realización de estas prácticas preventivas alguna vez en la vida, aunque éstas disminuyen al preguntar por el tiempo transcurrido desde la última revisión, al igual que en el estudio de Vacek et al³⁶.

La concordancia absoluta de la accidentalidad en los últimos 12 meses es también muy elevada, aunque el índice kappa es moderado. No hemos encontrado referencias donde se realice una pregunta similar. Sin embargo, Kozio-McLain et al³⁷ observan un alto grado de concordancia en la accidentalidad debida a accidentes de tráfico o por actos de violentos.

Respecto a la utilización del cinturón de seguridad, se obtienen altas concordancias absolutas, superiores en su utilización por carretera que por ciudad, posiblemente debido al menor grado de concienciación y su uso más infrecuente e irregular por ciudad. Stein et al^{21,25} observan también una alta concordancia.

Uno de los aspectos relevantes de este estudio es que ha sido realizado sobre una amplia muestra representativa de población adulta, cuya alta tasa de respuesta minimiza la posibilidad de haber introducido sesgos de selección. Por otro lado, se han evaluado las

preguntas originales del cuestionario, manteniendo las categorías de respuesta sin agrupar. En este tipo de estudios es habitual estimar la reproducibilidad de los indicadores de salud obtenidos de la categorización de la variable original, reduciendo a 2 el número de categorías, con lo que se incrementa artificialmente el valor de la concordancia.

El intervalo transcurrido entre la primera y la segunda entrevistas fue diseñado para que no fuera demasiado breve —con lo que se minimizaba el recuerdo de los entrevistados sobre las respuestas dadas en la primera entrevista— y no lo suficientemente largo como para que pudieran producirse cambios reales del comportamiento. En el análisis de discordancia, no se ha detectado un efecto del tiempo transcurrido entre las 2 entrevistas, al igual que se ha observado en otros estudios^{23,25}. La participación en la primera entrevista parece que no ha repercutido en la modificación de los comportamientos o prácticas preventivas, ya que las prevalencias estimadas en la segunda entrevista, además de que son muy parecidas, no han variado en una dirección que haga sospechar un cambio en los estilos de vida o un sesgo de recuerdo.

El test de la kappa ha sido ampliamente recomendado y utilizado como medida de concordancia. Sin embargo, su dependencia de la prevalencia ha sido reconocido como un efecto indeseable³⁸. Cuando la prevalencia tiende a 0 o a 100, así como cuando los marginales están simétricamente desequilibrados, se producen valores bajos del índice kappa³⁹. Otro factor que debe ser considerado en la interpretación de los coeficientes kappa es el número de categorías, ya que sus valores no ponderados disminuyen con el aumento de categorías⁴⁰. Sin embargo, mientras unos autores recomiendan la utilización del coeficiente kappa ponderado en datos ordinales⁴¹, otros critican este método ya que pueden conseguirse valores altos cuando la proporción de acuerdo exacto es baja¹⁸. Algunos ejemplos de estas limitaciones en nuestros datos podemos observarlos en la gran diferencia entre la concordancia

observada y los coeficientes kappa en variables como la medición de la PA alguna vez en la vida o la no utilización del cinturón de seguridad por carretera, cuyas concordancias observadas superan el 95% y las relativas no alcanzan valores de kappa de 0,5. Por otro lado, la concordancia exacta observada en el consumo de alcohol en el último mes, con 7 categorías de respuesta, es del 65,1%, y el valor del coeficiente kappa no ponderado de 0,572 se incrementa hasta 0,840 cuando se pondera, lo que demuestra una buena concordancia parcial. Es decir, aunque la concordancia exacta no sea muy alta, no se produce una gran desviación entre categorías.

Sin embargo, un alto nivel de reproducibilidad no significa que el cuestionario tenga alta validez de criterio. Desafortunadamente, no disponemos por el momento en el SIVFRENT de una comparación con parámetros objetivos que sirvan de patrón de referencia para valorarlo. Los estudios de validación del cuestionario del Behavioral Risk Factor Surveillance System concluyen que tiene una alta validez para las variables de consumo de tabaco, cribado de la PA e índice de masa corporal, moderada para el tiempo de realización de la última mamografía y sedentarismo en tiempo libre, y bajas para la estimación de PA o colesterol elevados^{33,42}.

En conclusión, el cuestionario empleado en el SIVFRENT obtiene estimaciones consistentemente elevadas en la prevalencia por grupos cuando se aplica repetidamente a los mismos individuos, y también describe una buena concordancia individual. No se observa un patrón uniforme de asociación entre las variables sociodemográficas y la discordancia.

Agradecimientos

A Ana Gandarillas, por la revisión de versiones previas de este manuscrito.

Bibliografía

1. Área de análisis epidemiológico y situación de salud. Centro Nacional de Epidemiología. Instituto de Salud Carlos III. Disponible en: http://193.146.50.130/mortal/mortal2000/w9100_cau.htm
2. Mc Ginnis JM, Foege WH. Actual causes of death in the United States. *JAMA* 1993;270:2207-12.
3. Hunink MG, Goldman L, Tosteson AN, et al. The recent decline in mortality from coronary heart disease, 1980-1990. The effect of secular trends in risk factors and treatment. *JAMA* 1997;277:535-42.
4. Kuulasmaa K, Tunstall-Pedoe H, Dobson A, et al. Estimation of contributing of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations. *Lancet* 2000;355:675-87.
5. McKean-Cowdin R, Feigelson HS, Ross RK, et al. Declining cancer rates in the 1990s. *J Clin Oncol* 2000;18:2258-68.
6. Dirección General de Planificación, Formación e Investigación. Plan de Salud de la Comunidad de Madrid. Madrid: Consejería de Salud, 1995.
7. Consejería de Sanidad y Servicios Sociales. Sistema de vigilancia de factores de riesgo asociados a enfermedades no transmisibles (SIVFRENT). *Boletín epidemiológico de la Comunidad de Madrid* 1996;4:3-15.
8. Centers for Disease Control and Prevention. Public health surveillance for behavioral risk factors in a changing environment: recommendations from the Behavioral Risk Factor Surveillance Team. *MMWR. Morb Mortal Wkly Rep* 2003; 52:1-11.
9. Marcus AC, Crane L. Telephone surveys in public health research. *Med Care* 1986;24:97-112.

10. Encuesta a hogares españoles sobre equipamiento, uso y valoración de servicios de telecomunicaciones, audiovisual e Internet. Informe preliminar (diciembre de 2002). Comisión del mercado de las telecomunicaciones e Instituto Nacional de Estadística. Disponible en: <http://www.aece.org/docs/EstudioINE+CMT.2002.pdf>
11. Nicholls II WL. Computer-assisted telephone interviewing: a general introduction. En: Groves RM, Biemer PP, Lyberg LE, Massey JT, Nicholls II WL, Waksberg J, editors. Telephone survey methodology. New York: John Wiley & Sons, 1988; p. 377-85.
12. Groves RM, Mathiowetz NA. Computer assisted telephone interviewing: effects on interviewers and respondents. Public Opinion Q 1984;48:356-69.
13. Ainsworth BE, Haskell VI, Leon AS, et al. Compendium of physical activities: classification of energy cost of human physical activities. Med Sci Sports Exerc 1993;25:71-80.
14. Armstrong BK, White E, Saracci R. Validity and reliability studies. En: Principles of exposure measurement in epidemiology. New York: Oxford University Press, 1993; p. 78-114.
15. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley & Sons, 1981.
16. Landis JR, Koch GC. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
17. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ Psychol Meas 1973;33:613-9.
18. Graham P, Jackson R. The analysis of ordinal agreement data: beyond weighted Kappa. J Clin Epidemiol 1993;46:1055-62.
19. Álvarez C, Alonso J, Domingo A, Regidor E. La medición de la clase social en ciencias de la salud. Informe de un grupo de trabajo de la Sociedad Española de Epidemiología. Sociedad Española de Epidemiología. Barcelona: SG editores, 1995.
20. Coughlin SS, Pickle LW, Goodman MT, Wilkens LR. The logistic modeling of interobserver agreement. J Clin Epidemiol 1992;45:1237-41.
21. Stein AD, Lederman RI, Shea S. The Behavioral Risk Factor Surveillance System questionnaire: its reliability in a statewide sample. Am J Public Health 1993;83:1768-72.
22. Starr GJ, Dal Grande E, Taylor AW, Wilson DH. Reliability of self-reported behavioural health risk factors in a South Australian telephone survey. Aust N Z Public Health 1999;23:528-30.
23. Browson RC, Jackson-Thompson J, Wilkerson JC, Kiani F. Reliability of information on chronic disease risk factors collected in the Missouri Behavioral Risk Factor Surveillance System. Epidemiology 1994;5:545-9.
24. Shea S, Stein AD, Lantigua R, Basch CE. Reliability of the Behavioral Risk Factor Survey in a triethnic population. Am J Epidemiol 1991;133:489-500.
25. Stein AD, Courval JM, Lederman RI, Shea S. Reproducibility of responses to telephone interviews: demographic predictors of discordance in risk factor status. Am J Epidemiol 1995;141:1097-106.
26. Martkainen P, Aromaa A, Heliövaara M, Klaukka T, Knekt P, Maatela J, et al. Reliability of perceived health by sex and age. Soc Sci Med 1999;48:1117-22.
27. Lundberg O, Manderbacka K. Assessing reliability of a measure of self-rated health. Scand J Soc Med 1996;24:218-24.
28. Kumpusalo E, Pekkarinen H, Neittaanmäki L, Penttilä I, Halonen P. Identification of health status dimensions in a working-age population: an exploratory study. Med Care 1992;30:392-9.
29. Centers for Disease Control and Prevention. Monthly estimates of leisure-time physical inactivity: United States, 1994. MMWR. Morb Mortal Wkly Rep 1997;46:393-7.
30. Pols MA, Peeters PHM, Ocké MC, Slimani N, Bueno-de-Mesquita HB, Collette HJA. Estimation of reproducibility and relative validity of the questions included in the EPIC physical activity questionnaire. Int J Epidemiol 1997;26:181S-9S.
31. Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. Validity and repeatability of the EPIC-Norfolk physical activity questionnaire. Int J Epidemiol 2002;31:168-74.
32. Batty D. Reliability of a physical activity questionnaire in middle-aged men. Public Health 2000;114:474-6.
33. Bowlin SJ, Morrill Bd, Nafziger AN, Lewis C, Pearson TA. Reliability and changes in validity of self-reported cardiovascular disease risk factors using dual response: the Behavioral Risk Factor Survey. J Clin Epidemiol 1996;49:511-7.
34. Subías-Loren PJ, Casanovas-Cuquet E. Factores asociados con una mejor cumplimentación del examen periódico de salud en el adulto. Aten Primaria 1998;22:570-3.
35. Stein AD, Lederman RI, Shea S. Reproducibility of the women's module of the Behavioral Risk Factor Surveillance System questionnaire. Ann Epidemiol 1996;6:47-52.
36. Vacek PM, Mickey RM, Worden JK. Reliability of self-reported breast screening information in a survey of lower income women. Prev Med 1997;26:287-91.
37. Koziol-McLain J, Brand D, Morgan D, Leff M, Lowenstein R. Measuring injury risk factors: question reliability in a statewide sample. Injury Prev 2000;6:148-50.
38. Chinn S, Burney PGJ. On measuring repeatability of data from self-administered questionnaires. Int J Epidemiol 1987;16:121-7.
39. Feinstein AR, Cicchetti DV. High agreement but low kappa (I). The problems of two paradoxes. J Clin Epidemiol 1990;43:543-9.
40. Brenner H, Kliebsch U. Dependence of weighted Kappa coefficients on the number of categories. Epidemiology 1996;7:199-202.
41. Maclure M, Willett WC. Misinterpretation and misuse of the Kappa statistic. Am J Epidemiol 1987;126:161-9.
42. Nelson DE, Holtzman D, Bolen J, Stanwyck CA, Mack KA. Reliability and validity of measures from the Behavioral Risk Factor Surveillance System (BRFSS). Soz Praventivmed 2001;46:3S-42S.