

## CRITERIOS PARA DEFINIR ALCANCE DE CONSENSO EN ESTUDIOS DELPHI QUE VALORAN MEDICAMENTOS POTENCIALMENTE INAPROPIADOS EN ADULTOS MAYORES

*CRITERIA FOR DEFINING CONSENSUS ACHIEVEMENT IN DELPHI STUDIES THAT ASSESS POTENTIALLY INAPPROPRIATE MEDICATIONS IN THE ELDERLY*

Marta Marzi <sup>1</sup>, Miryam Pires <sup>2</sup>, Nora Quaglia <sup>2</sup>

---

### Resumen

**Introducción:** El tratamiento estadístico de respuestas a encuestas basadas en una escala categórica tipo Likert, como las empleadas en la técnica Delphi, ha sido cuestionado por utilizarse metodología basada en el análisis de datos numéricos.

**Objetivo:** Desarrollar criterios para definir el alcance de consenso en estudios Delphi realizados para la valoración de medicamentos potencialmente inapropiados (MPI) en adultos mayores.

**Métodos:** Se construyó un indicador ( $Y_q$ ) que contempla el acuerdo por pares de evaluadores, cuya fórmula de cálculo se basa en el análisis combinatorio.  $Y_q$  se aplicó en un estudio piloto con metodología Delphi para la valoración de la seguridad de 12 medicamentos a través de una escala Likert con cinco categorías de respuestas.

**Resultados:** En base al análisis de la distancia ( $d$ ) entre cada par de categorías de la escala se determinó la ponderación  $w=1-d$  asociada, para ser aplicada en el cálculo de  $Y_q$ : proporción de acuerdos ponderados. Se definieron tres criterios que debían ser satisfechos para el alcance de consenso en cada ítem (medicamento) del cuestionario Delphi: a) Cantidad de evaluadores  $\geq 60\%$  de los integrantes del panel, b)  $Y_q \geq 0,800$ ; c) Frecuencia de la moda  $\geq 60\%$ . Según este análisis 8/12 medicamentos fueron valorados como potencialmente inapropiados para uso en adultos mayores mientras que los 4 restantes debieron ser reevaluados en rondas sucesivas por no lograrse el consenso.

**Conclusión:** El indicador contempla las distancias reales entre las categorías de la escala Likert y los criterios desarrollados constituyen una herramienta simple para el análisis de las encuestas Delphi realizadas para la valoración de MPI en adultos mayores.

**Palabras claves:** Consenso, Método Delphi, Prescripción Inapropiada, Adultos Mayores, Escala Likert.

### Abstract

**Background:** The statistical analysis of survey responses based on a categorical Likert scale, as those used in the Delphi technique, has been questioned because the methodology used is based on numerical data analysis.

---

<sup>1</sup>Departamento de Matemática y Estadística, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario. Rosario, Argentina.

<sup>2</sup>Centro de Tecnología en Salud Pública, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario. Rosario, Argentina.

Correspondencia:

Marta Mónica Marzi.

Dirección postal: Presidente Julio Roca, N° 653, Piso 7, Dto.A. Rosario, (Santa Fe), Argentina. Teléfono: +54 341 4253300. Dirección electrónica: marzimarta58@gmail.com

Fuentes de apoyo financiero: El proyecto que dio origen a este trabajo no tuvo apoyo financiero.

**Aim:** To develop criteria for defining the consensus achievement in Delphi studies conducted for the assessment of potentially inappropriate medications (PIM) in the elderly.

**Methods:** It was constructed an index ( $Y_q$ ) which takes into account the agreement by pairs of evaluators, whose calculus equation is based on combinatorial analysis.  $Y_q$  was applied in a pilot study with Delphi methodology for assessing the safety of 12 drugs through a Likert scale with five response categories.

**Results:** On the basis of analyzing the distance ( $d$ ) in between each pair of categories from the scale, it was determined the associated weighting,  $w=1-d$ , to be applied in the calculus of  $Y_q$ : proportion of agreements weighted. There were defined three criteria that need to be satisfied to obtain the consensus in each item (drug) of Delphi questionnaire: a) Number of evaluators  $\geq 60\%$  of the panel members, b)  $Y_q \geq 0,800$ ; c) frequency of the statistical mode  $\geq 60\%$ . On regards to the analysis, 8/12 drugs were evaluated as potentially inappropriate for elderly adults while the remaining 4 should be revalued in successive rounds due to not obtaining consensus.

**Conclusion:** The index takes into account the real distances between the Likert scale categories and the developed criteria constitute a simple tool for the analysis of the Delphi questionnaires made for the valuation of PIM on older adults.

**Key Words:** Consensus, Delphi Study, Inappropriate Prescribing, Older Adults, Likert Scale.

## Introducción

La técnica Delphi, es un método ampliamente utilizado y aceptado para lograr la convergencia de opiniones de un grupo de expertos sobre temas de los que el conocimiento existente es incierto o incompleto. Es un método relativamente simple y económico que permite explorar sistemáticamente el conocimiento y experiencia de personas calificadas en el tema problema a través de cuestionarios. El método Delphi se basa en un proceso iterativo que alienta a los participantes seleccionados a reevaluar sus juicios iniciales sobre la información proporcionada en iteraciones anteriores. Una de las características principales es el anonimato entre los participantes ya que las consultas y devoluciones de las respuestas se realizan por correo, evitando así posibles fuentes de sesgos<sup>1,2,3</sup>. La técnica Delphi se ha aplicado con múltiples adaptaciones en su aplicación y análisis de los resultados<sup>4,5,6</sup>. En el ámbito de la salud, en los últimos veinte años, se están elaborando listas de Medicamentos Potencialmente Inapropiados (MPI) para uso en adultos mayores mediante la consulta a un panel de expertos aplicando la técnica Delphi<sup>7,8</sup>. Básicamente el cuestionario consiste en una lista de medicamentos con información acerca del potencial riesgo de efectos adversos

en ancianos. Para la valoración de cada medicamento se emplea generalmente una escala ordinal tipo Likert con cuatro o cinco alternativas puntuadas<sup>7-16</sup>.

Una escala Likert se emplea habitualmente cuando se diseñan cuestionarios para recabar valoraciones de aspectos que no son medibles en forma exacta<sup>17</sup>. Con ella se pretende conocer la actitud de los encuestados ante un tema determinado. Básicamente, la encuesta consiste en una serie de ítems o enunciados ante los cuales los individuos tienen que manifestar su grado de acuerdo o desacuerdo eligiendo una de  $k$  categorías de la escala. El número  $k$  de categorías varía normalmente entre 3 y 10, aunque la elección más conveniente desde un punto psicométrico y estadístico es entre 4 y 7<sup>18</sup>.

Las principales estadísticas utilizadas para determinar el arribo al consenso de los expertos en la valoración de los MPI son medidas de tendencia central (media, mediana y moda) y medidas de dispersión (desviación estándar y rango intercuartil). Sin embargo, la conveniencia de utilizar la media es cuestionada si las escalas utilizadas no están diseñadas a intervalos iguales, como ocurre generalmente en una escala tipo Likert<sup>19</sup>. El uso del rango inter-cuartil también ha sido cri-

ticado como inadecuado por no aportar información acerca de la distribución de las respuestas<sup>20</sup>. El tratamiento estadístico de respuestas basadas en una escala Likert consistiría en principio en un estudio de datos categóricos para cuyo análisis debería recurrirse a gráfico de barras o de sectores. Sin embargo, se suelen codificar tales datos categóricos mediante números enteros y aplicar métodos estadísticos para datos numéricos. Ambos tratamientos resultan bastante limitados puesto que el número de valores distintos es muy pequeño; además, una codificación entera no suele reflejar las diferencias reales entre las distintas categorías de la escala ya que es abrupta y difícilmente puede captar la imprecisión subyacente a las respuestas<sup>19</sup>.

Por otra parte, cuando se trata de analizar la concordancia entre observadores utilizando variables categóricas para una muestra, es usual la implementación del índice kappa el cual, básicamente, relaciona las proporciones de acuerdos encontrados entre observadores con las esperadas por azar<sup>21,22</sup>. Sin embargo, este índice y sus variantes resulta útil para el estudio del conjunto muestral, mientras que en el caso de los MPI, interesa sistematizar la valoración para cada ítem (medicamento), separadamente del resto.

El objetivo del presente trabajo es desarrollar criterios para definir el alcance de consenso en estudios que emplean el método Delphi para la valoración, basada en una escala ordinal tipo Likert, de medicamentos que usados en los adultos mayores podrían ser inapropiados.

### Métodos

Se realizó un estudio piloto aplicando el método Delphi para valorar una lista de 12 medicamentos ampliamente comercializados en Argentina, que fueron propuestos en algunos países como "potencialmente inapropiados" (PI) para uso en los adultos mayores. Se contactó vía e-mail a profesionales de la salud con reconocida idoneidad en el tema a consultar, solicitando su participación en el estudio en calidad de panelistas, previa información acerca del método Delphi.

La Encuesta consistió en completar una tabla en la que se listaron los medicamentos PI. Por cada fármaco propuesto había un espacio en blanco que el panelista debía completar con la valoración que, según una escala Likert, asignaba a la siguiente afirmación:

MEDICAMENTO POTENCIALMENTE INAPROPIADO para uso en adultos mayores independiente del diagnóstico o condición clínica del paciente".

La escala Likert contenía 4 categorías ordenadas según el nivel decreciente de acuerdo con la afirmación enunciada: A, B, D, E. Se incluyó una quinta categoría para contemplar la posibilidad de abstención y evitar que el espacio destinado a la valoración quedara en blanco. A esta categoría se le asignó la letra C para que quedara ubicada en la posición central de la escala, marcando la diferencia entre las categorías A y B que equivalen a la inclusión del fármaco como PI, de las categorías D y E que lo excluyen.

### Escala Likert con el detalle de las valoraciones utilizadas:

A. Muy de acuerdo. El uso de este medicamento debe evitarse pues hay alternativas terapéuticas más seguras.

B. De acuerdo. Debería evitarse pero no hay alternativas terapéuticas más seguras. Sugiero administrarlo con precaución.

C. Neutral. No tengo opinión sobre este medicamento.

D. En desacuerdo. El riesgo de uso en ancianos es comparable con el riesgo en adultos jóvenes.

E. Totalmente en desacuerdo. Medicamento seguro.

Por cada medicamento listado había además tres espacios adicionales que el encuestado debía completar, con alternativas terapéuticas más seguras si su valoración fuera "A", con recomendaciones/precauciones si su valoración fuera "B" y con observaciones si lo considerara pertinente.

### Determinación del acuerdo por pares

Se construyó el indicador "proporción de acuerdos ponderados" ( $Y_q$ ) que contempla el acuerdo por pares de evaluadores:

$$Y_q = \frac{\sum_{i=1}^k \sum_{j=1}^k w_{ij} n_i n_j - N_q}{N_q \cdot (N_q - 1)}$$

Siendo  $q$  el medicamento de la lista clasificado en una de las  $k$  categorías de la escala Likert por cada uno de los  $N_q$  panelistas que lo evaluaron. El número de panelistas que clasificaron el medicamento  $q$  en la  $k$ -ésima categoría es  $n_k$  y la ponderación correspondiente al acuerdo-desacuerdo entre las categorías  $i$  y  $j$  es  $w_{ij}$ , que

cumple las siguientes condiciones: . En la Figura 1 se muestra el desarrollo de la fórmula de cálculo de  $Y_q$ .

La ponderación  $w$  entre los pares de categorías de la escala Likert se asignó teniendo en cuenta la distancia entre las mismas como indicador de disimilitud.

En general se da el nombre de distancia entre dos puntos  $i$  y  $j$  (categorías, objetos, individuos) a una medida, indicada por  $d(i,j)$ , que mide el nivel de semejanza entre ambos puntos, en relación a un cierto número de características cuantitativas y/o cualitativas.

Deben verificarse las siguientes propiedades:  $d(i,j) > 0$  (no negatividad);  $d(i,i) = 0$ ;  $d(i,j) = d(j,i)$  (simetría). Si los puntos están alineados se verifica además:  $d(i,j) = d(i,t) + d(t,j)$ , siendo  $t$  un punto entre  $i$  y  $j$ .

Se consideró que las categorías de la escala Likert se hallan alineadas, de modo que, fijada la distancia máxima  $d_{max}$  entre las categorías extremas, la ponderación correspondiente al acuerdo entre las categorías  $i$  y  $j$  se obtuvo calculando la diferencia entre  $d_{max}$  y la distancia entre las mismas:  $w_{ij} = d_{max} - d(i,j)$ .

El procedimiento empleado para la elección del punto de corte del indicador  $Y_q$  consistió en:

\* Calcular el valor más bajo que podría asumir el indicador ( $Y_q^*$ ) cuando el número de panelistas que evalúan un medicamento es un mínimo prefijado como aceptable ( $N_q^*$ ).

\* Promediar  $Y_q^*$  con el máximo valor que puede asumir el indicador ( $Y_q = 1$ ).

\* Sumar 0,100 unidades al promedio obtenido para asegurar la separación entre alcanzado/no alcanzado el acuerdo por pares.

$$\frac{1 + Y_q^*}{2} + 0,100$$

Para resumir la valoración según la escala Likert de los medicamentos en los que se logró el acuerdo por pares, se eligió la moda como medida de tendencia central.

Análisis del nivel de consenso

Para la formulación de los criterios que definirían el alcance de consenso en cada ítem de la encuesta se contempló: el número de evaluadores, el valor de  $Y_q$  y la frecuencia de cada valoración. Se utilizó el programa Microsoft Excel 2007 para integrar toda la información de las encuestas, con la planilla de cálculo se obtuvo la moda y

también  $Y_q$  programando la fórmula [1] del desarrollo que se muestra en la Figura 1.

Acuerdo por pares

Sea

$q$ : medicamento de la lista clasificado por cada evaluador en una categoría de la escala Likert

$k$ : número de categorías ordinales de la escala

$M$ : número de expertos que integran el panel

$N_q$ : número de panelistas que valoraron el medicamento  $q$  ( $N_q \leq M$ )

$n_i$ : número de evaluadores que clasifican el medicamento  $q$  en la  $i$ -ésima categoría

$w_{ij}$ : ponderación correspondiente al acuerdo-desacuerdo entre las categorías  $i$  y  $j$ , con las condiciones:

$$w_{ii} = 1 ; 0 \leq w_{ij} < 1 \quad \forall i \neq j ; w_{ij} = w_{ji}$$

Para el medicamento  $q$ , el número de acuerdos ponderados es:

$$X_q = \sum_{i=1}^k w_{ii} C(n_i, 2) + \sum_{i=1}^k \sum_{j>i}^k w_{ij} n_i n_j$$

siendo  $C(n_i, 2)$  las combinaciones de  $n_i$  elementos formando subgrupos de 2.

Como el número de posibles pares de clasificación para el medicamento  $q$  es  $C(N_q, 2)$ , la

proporción de acuerdos ponderados para dicho medicamento es:

$$Y_q = \frac{X_q}{C(N_q, 2)} = \frac{\sum_{i=1}^k w_{ii} C(n_i, 2) + \sum_{i=1}^k \sum_{j>i}^k w_{ij} n_i n_j}{C(N_q, 2)} = \frac{\frac{1}{2} \sum_{i=1}^k w_{ii} n_i (n_i - 1) + \sum_{i=1}^k \sum_{j>i}^k w_{ij} n_i n_j}{\frac{N_q (N_q - 1)}{2}} \quad [1]$$

$$= \frac{\sum_{i=1}^k w_{ii} n_i (n_i - 1) + 2 \sum_{i=1}^k \sum_{j>i}^k w_{ij} n_i n_j}{N_q (N_q - 1)} = \frac{\sum_{i=1}^k \sum_{j=1}^k w_{ij} n_i n_j - N_q}{N_q (N_q - 1)} \quad [2]$$

$$\text{dado que: } \sum_{i=1}^k w_{ij} n_i = \sum_{i=1}^k n_i = N_q$$

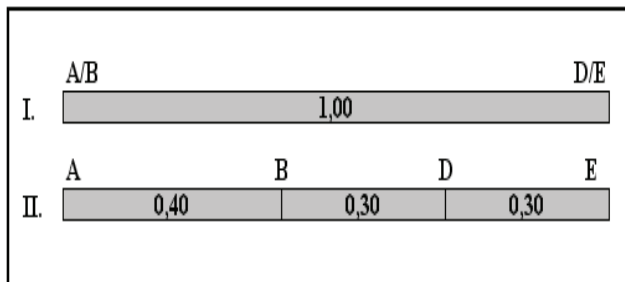
Nota: La categoría C no está contemplada dentro de la numeración  $k$  dado que corresponde a las abstenciones.

## Resultados

El panel quedó constituido por diez expertos, a quienes se les envió la encuesta de la primera Ronda Delphi con las instrucciones necesarias para completarla y las razones por las cuales los medicamentos listados fueron considerados PI en estudios previos 7-16.

Asignación de la ponderación  $w$  a cada par de categorías  $k$

En primer lugar se determinó la distancia  $d(i,j)$  entre cada par de categorías de la escala fijando la distancia máxima en  $d_{max}=1$ , para después calcular la ponderación  $w(i,j)$  correspondiente. La cuantificación de la distancia se realizó atendiendo al enunciado consultado en la encuesta y a las categorías propuestas en la escala. Se contemplaron dos enfoques diferentes (Figura 2):



I. El medicamento. Acuerdo/Desacuerdo en que el uso del medicamento supone un incremento del riesgo de efectos adversos en los adultos mayores

II. La prescripción. No prescripción/Prescripción del medicamento atendiendo a su nivel de seguridad en relación a la edad del paciente. Según el primer enfoque, no habría distancia (distancia 0) entre las categorías A y B por un lado y las categorías D y E por el otro. En cambio entre las categorías A, B y D, E la distancia se consideró máxima e igual a 1 (Tabla 1).

Tabla 1

	A	B	D	E
A	0	0	1	1
B	0	0	1	1
D	1	1	0	0
E	1	1	0	0

Según el segundo enfoque (Tabla 2) y recordando que la valoración A indica que debe evitarse la prescripción del fármaco pues hay alternativas terapéuticas más seguras, se estableció:

Tabla 2

	A	B	D	E
A	0	0,4	0,7	1
B	0,4	0	0,3	0,6
D	0,7	0,3	0	0,3
E	1	0,6	0,3	0

Distancia entre A y B: 0,40. La valoración B indica que el fármaco debe prescribirse con precaución en los adultos mayores

Distancia entre A y D: 0,70. La valoración D indica que se prescribirá atendiendo a las precauciones habituales, sin reconocimiento de riesgo incrementado por la edad.

Distancia entre A y E: 1. Los fármacos de la categoría E serán prescritos como todo medicamento considerado seguro.

Por otra parte, y dado que la categoría B asume la prescripción con precaución para los adultos mayores:

Distancia entre B y D: 0,30. La categoría D admite prescripción sin atención a riesgo incrementado por la edad.

Distancia entre B y E: 0,60. La categoría E lo considera un medicamento seguro para cualquier población.

Por último:

Distancia entre D y E: 0,30. Los medicamentos de la categoría D y E se prescribirán en el adulto mayor igual que en los jóvenes, aunque la valoración D reconoce que el medicamento puede no ser seguro.

Finalmente se obtuvo un único valor  $d(i,j)$  entre cada par de categorías promediando las distancias asignadas según los dos enfoques contemplados separadamente (Tabla 3).

Tabla 3

	A	B	D	E
A	0	0,2	0,85	1
B	0,2	0	0,65	0,8
D	0,85	0,65	0	0,15
E	1	0,8	0,15	0

La ponderación  $w(i,j)$  resultó de restar de la unidad cada valor de distancia. Quedó así conformada una tabla de doble entrada con las ponderaciones correspondientes a cada par de categorías, para ser utilizada en el cálculo de  $Y_q$  (Tabla 4).

Tabla 4

	A	B	D	E
A	1	0,8	0,15	0
B	0,8	1	0,35	0,2
D	0,15	0,35	1	0,85
E	0	0,2	0,85	1



Se consideró que hubo acuerdo entre los evaluadores en aquellos medicamentos con una proporción de acuerdos ponderados mayor o igual que 0,800 ( $Y_q \geq 0,800$ ); este punto de corte se obtuvo conforme a lo explicitado en "Metodos". Formulación de los criterios para definir el alcance de consenso

### Se adoptaron tres criterios:

a- Cantidad de evaluadores ( $N_q$ ): no inferir al 60% de los integrantes del panel

b- Proporción de acuerdos ponderados: mayor o igual que 0,800.

c- Moda: frecuencia no inferior al 60%.

Se consideró alcanzado el consenso si eran satisfechos los tres criterios. Aquel medicamento con al menos un criterio no satisfecho sería reevaluado en una 2ª Ronda Delphi.

### Resultados de la 1ª Ronda Delphi

Se alcanzó el consenso en 8 de los 12 fármacos evaluados (categoría A: 5, categoría B: 3). De los 4 fármacos en los que no se arribó a consenso, en 3 no se satisfizo uno de los tres criterios adoptados (criterio "a": rilmenidina, "b": meloxicam, "c": haloperidol), y en 1 fueron dos los criterios no satisfechos (criterios "b" y "c": dipiridamol). En la Tabla 5 se presenta, a modo de ejemplo, 6 de los medicamentos incluidos en la encuesta, con la valoración de los panelistas basada en la escala Likert y los resultados del análisis del nivel de consenso alcanzado atendiendo a los criterios adoptados.

Como puede observarse, por ejemplo para indometacina, uno de los 10 panelistas le asignó la categoría C (absención), resultando en 9 el número real de evaluadores para este medicamento (90%) y satisfecho el criterio "a" pues supera el mínimo exigido.

Para obtener la proporción de acuerdos ponderados se programó en Excel la fórmula [1]. Se

mostrará el cálculo de  $Y_q$  realizado en forma manual para indometacina, asignando a las categorías A, B, D, E los números 1, 2, 3 y 4 respectivamente. Se tiene que  $n_1 = 6$ ,  $n_2 = 3$ ,  $n_3 = 0$ ,  $n_4 = 0$ ,  $N_q = 9$  y consultando la tabla de las ponderaciones resulta:

Según la fórmula (1):

$$Y_q = \frac{1 \cdot \frac{6 \cdot 5}{2} + 1 \cdot \frac{3 \cdot 2}{2} + 0,80 \cdot 6 \cdot 3}{\frac{9 \cdot 8}{2}} = 0,900$$

Según la fórmula (2):

$$Y_q = \frac{1 \cdot 6 \cdot 6 + 1 \cdot 3 \cdot 3 + 0,80 \cdot 6 \cdot 3 + 0,80 \cdot 6 \cdot 3 - 9}{9 \cdot 8} = 0,900$$

Dado que  $Y_q \geq 0,800$ , el criterio "b" también fue satisfecho. Por último, la categoría más seleccionada resultó ser la "A" con una frecuencia del 67% (6/9) no inferior al 60% según el criterio "c". Por lo tanto se consideró alcanzado el consenso para incluir a indometacina en el grupo de los MPI cuyo uso debe evitarse en los adultos mayores por existir alternativas terapéuticas más seguras.

Los items en los que no se logró consenso fueron reevaluados en rondas sucesivas adoptándose, para el análisis de los resultados, el mismo criterio que para la encuesta de la 1ª Ronda Delphi.

### Discusión

El método Delphi es un recurso valioso en la investigación de MPI, sin embargo el manejo de los datos obtenidos de las encuestas plantea varias dificultades. En este trabajo se proponen tres criterios para definir el alcance de consenso, el cual es innovador en relación a trabajos previos publicados que emplearon el método Delphi para la valoración de MPI.

Se construyó el indicador  $Y_q$ , que contempla las distancias reales entre las diferentes categorías de la escala Likert según dos enfoques claramente definidos, el medicamento y la prescrip-

Tabla 5

Medicamento	Valoración de cada panelista										Criterios			Resultados 1ª Ronda
	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	a	b	c	
Indometacina	A	C	A	B	A	B	A	A	A	B	90%	0,9	A 6/9=67%	A
Amiodarona	C	C	A	B	A	B	B	B	B	B	80%	0,914	B 6/8=75%	B
Rilmenidina	C	C	A	A	C	C	A	A	C	B	50%	0,92	A 4/5=80%	Sin consenso
Meloxicam	A	C	A	A	A	B	A	E	A	B	90%	0,722	A 6/9=67%	Sin consenso
Haloperidol	B	B	B	A	A	B	A	A	A	B	100%	0,889	A/B 5/10=50%	Sin consenso
Dipiridamol	A	B	A	A	C	D	A	D	A	B	90%	0,636	A 5/9=55%	Sin consenso

ción. El cálculo de  $Y_q$ , es una alternativa al uso de métodos estadísticos para el manejo de datos numéricos, cuestionado por discretizar variables categóricas en un número reducido de valores distintos y ocultando además la subjetividad implícita en las valoraciones<sup>19</sup>.

La fórmula desarrollada para el cálculo de  $Y_p$  es sencilla pues se basa en el Análisis Combinatorio. También su programación es sencilla y puede hacerse con un conocimiento mínimo de las herramientas que ofrece la planilla de cálculo de cualquier versión del programa Microsoft Excel. También a diferencia de trabajos previos, se impuso un número mínimo de evaluadores en relación al total de panelistas como uno de los criterios para determinar el consenso, lo que confiere mayor fiabilidad a los resultados.

Este trabajo ofrece una herramienta simple para el análisis de las respuestas a encuestas que utilizan escalas categóricas tipo Likert cuando se implementa el método Delphi para la elaboración de listas de MPI en adultos mayores.

### Agradecimientos

Los autores agraden a los expertos que participaron en las Rondas Delphi: Dr Guillermo Di Girolamo (Médico Esp Farmacología, Fac Medic Univ de Buenos Aires), Dr Mariano Nuñez (Médico Esp Medic Interna, Fac Medic Univ de Buenos Aires), Dr Juan Martini (Médico Esp Geriátrica y Gerontología, Hosp Geriátrico Provincial de Rosario), Dr José María Puig (Médico Esp Cardiología, Fac Cs Médicas, Univ Nac de Rosario), Dr Ignacio Demey (Médico Esp Neurología, Mg en Inv Clínica Farmacológica, FLENI, Escobar), Dr Pedro Lipszyc (Médico Esp Farmacología, Fac Medic Univ de Buenos Aires), Dr Sebastián Sevilla (Médico Esp Medic Interna, Hosp. Univ Austral, Buenos Aires), Dr Guido Iantorno (Médico Esp Gastroenterología y Clínica Médica, Hosp Udaondo, Buenos Aires), Dr Marcelo Loyato (Médico Esp Clínica Médica, Hosp Fernandez, Buenos Aires), Dr Roberto Diez (Médico, Prof Titular Cátedra Farmacología, Fac Medic, Universidad de Buenos Aires).

### Referencia

1. Darkley NC, Helmer O. An experimental application of The Delphi method to the use of experts. *The Rand Corp.* 1962; Memorandum. RM-727/1-abridged. Disponible en: [http://www.rand.org/content/dam/rand/pubs/research\\_memoranda/2009/RM727.1.pdf](http://www.rand.org/content/dam/rand/pubs/research_memoranda/2009/RM727.1.pdf) [Consultado el 1 de Julio de 2015]
2. Darkley NC. An experimental study of group opi-

nion. *The Rand Corp.* 1969; RM-5888-PR. Disponible en: [http://www.rand.org/content/dam/rand/pubs/research\\_memoranda/2005/RM5888.pdf](http://www.rand.org/content/dam/rand/pubs/research_memoranda/2005/RM5888.pdf) [Consultado el 1 de Julio de 2015]

3. Linstone HA, Turoff M. (1975). Introduction. En: Linstone HA & Turoff M, Editores, *The Delphi method: Techniques and Applications*. Addison-Wesley Publishing Company. Londres. Reproducción 2002. p. 3-12. Disponible en: <http://is.njit.edu/pubs/delphibook/delphibook.pdf> [Consultado el 15 de Julio de 2015]
4. Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *J Adv Nurs.* 2006; 53:205-12
5. Hasson F, Deeney S. Enhancing rigour in the Delphi technique research. *Technological Forecasting & Social Change* 2011; 78: 1695-704.
6. Vaidya SA, Manning SE, Dhankhar P, Meltzer M, Rupprecht Ch, Hull HF, Fishbein DB. Estimating the risk of rabies transmission to humans in the U.S: a Delphi analysis. *BMC Public Health* 2010; 10:278. Disponible en: <http://www.biomedcentral.com/1471-2458/10/278> [Consultado el 15 de Julio de 2015]
7. Beers MH. Explicit criteria for determining potentially inappropriate medications use by the elderly. *Arch Intern Med.* 1997; 157:1531-36.
8. McLeod PJ, Huang AR, Tamblyn RM, Gayton DC. Defining inappropriate practices in prescribing for elderly people: a national consensus panel. *Can Med Assoc J.* 1997; 156:385-39
9. Fick DM, Cooper JW, Wade WE, Waller JL, Ross Maclean J, Beers MH. Updating the Beers Criteria for Potentially Inappropriate Medication Use in Older Adults- Results of a US Consensus Panel of Experts. *Arch Intern Med.* 2003; 163:2716-24.
10. Laroche ML, Charmes JP, Merle L. Potentially inappropriate medications in the elderly: a French consensus panel list. *Eur J Clin Pharmacol.* 2007; 63:725-31
11. Gallagher P, Ryan C, Byrne S, Kennedy J, O'Mahony D. STOPP (Screening Tool of Older Person's Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment). Consensus validation. *Int J Clin Pharmacol Ther.* 2008; 46:72-83.
12. Rognstad S, Brekke M, Fetveit A, Spigset O, Wyller TB, Straand J. The Norwegian General Practice (NORGE) criteria for assessing potentially inappropriate prescriptions to elderly patients. A modified Delphi study. *Scand J Prim Health Care* 2009; 27:153-9.
13. Kim D, Heo SI, Lee SH. Development of a List of Potentially Inappropriate Drugs for the Korean Elderly Using the Delphi Method. *Health Inform Res.* 2010; 16(4):231-52.
14. Holt S, Schmiedl S, Thürmann PA. (2010). Potentially Inappropriate Medications in the Elderly: The PRISCUS List. *Dtsch Arztebl Int.* 2010; 107: 543-51.
15. Mann E, Böhmendorfer B, Frühwald T, Roller-Wirnsberger RE, Dovjak P, Dückelmann-Hofer C, Fischer P,

- Rabady S, Iglseider B. Potentially inappropriate medication in geriatric patients: the Austrian consensus panel list. *Wien Klin Wochenschr.* 2012; 124:160-69.
16. AGS. American Geriatrics Society Updated Beers Criteria for Potentially Inappropriate Medication Use in Older Adults. *J Am Geriatr Soc.* 2012; 60: 616-31.
17. Likert, R. A technique for the measurement of attitudes. *Archives of Psychology.* 1932; 140: 1-50. Traducción al castellano en C. H. Wainerman (Comp.) (1976), *Escalas de medición en ciencias sociales*, Buenos Aires: Nueva visión. pp-199-260.
18. Lozano LM, García-Cueto E, Muniz J. Effect of the number of response categories on the reliability and validity of rating scales. *Methodology.* 2008; 4:73-9
19. De la Rosa de Súa, Sara. Análisis estadístico comparativo de tres escalas de valoración: Likert, fuzzy-Likert y fuzzy de respuesta libre. Trabajo final de Master en Modelización Matemática, Estadística y Computación. Universidad de Oviedo, España, 2012. Disponible en: [http://digibuo.uniovi.es/dspace/bitstream/10651/4145/3/TFM\\_SaradelaRosadeSúa.pdf](http://digibuo.uniovi.es/dspace/bitstream/10651/4145/3/TFM_SaradelaRosadeSúa.pdf) [Consultado el 15 de Julio de 2015]
20. Scheibe M, Skutsch M, Schofer J.(1975). Experiments in Delphi Methodology. En: Listone H.A &Turroff M, Editores, *The Delphi Methodology. Techniques and Applications*; Addison-Wesley Publishing Company. Londres. Reproducción 2002; p: 257-81. Disponible en: <http://is.njit.edu/pubs/delphibook/ch4c.pdf> [Consultado el 15 de Julio de 2015]
21. Fleiss JL, Levin B, Cho Paik M. *Statistical Methods for Rates and Proportions. Third Edition.* Editors: Bolding DJ, Cressie AC, Fisher NI, Johnstone IM, Kadane JB, Ryan LM, Scott DW, Smith AFM, Teugels JL. 2003. Wiley & Sons, Inc. Publications.
22. Abaira V, Perez de Vargas A. Generalization of the kappa coefficient for ordinal categorical data, multiple observers and incomplete designs. *Qüestiló.* 1999; 23(3):561-571.