

La agregación de datos en la medición de desigualdades e inequidades en la salud de las poblaciones

Xavier Metzger¹

RESUMEN

Objetivos. Comparar cómo diferentes grados de agregación de datos repercuten en los resultados de la medición de desigualdades e inequidades de salud en una población y examinar su idoneidad para estudios sobre el tema.

Métodos. A modo de ejemplo, se calcularon las medidas más frecuentemente utilizadas para cuantificar las desigualdades e inequidades reflejadas por la tasa de mortalidad infantil en Costa Rica entre 1973 y 1984.

Resultados. Según los resultados obtenidos, las medidas de desigualdad presentadas (a excepción de las que se basaron en modelos de regresión) no parecen ser sensibles al grado de agregación de los datos utilizados cuando las unidades objeto del estudio son grupos socioeconómicos. Por el contrario, cuando las unidades comparadas son zonas geográficas, mayores grados de desagregación de los datos redundan en medidas que indican la presencia de un mayor grado de desigualdad.

Conclusiones. Los resultados indican que algunas medidas generan valores muy dispares según el nivel de agregación utilizado, por lo que se demuestra la importancia de elegir tanto las medidas como la agregación que sean adecuadas a la luz de los objetivos de cada estudio.

Palabras clave

Inequidad, desigualdad, medición de inequidades, indicadores, situación sanitaria.

Desde los albores de los años ochenta, la medición de las desigualdades e inequidades en el estado de salud de las poblaciones se ha convertido en una de las prioridades de los organismos nacionales e internacionales dedicados a la salud pública (1), movidos por la necesidad de conocer la situación de salud, tomar medidas para controlar sus tendencias a corto y largo plazo, y buscar medidas del impacto de las políticas sanitarias que indiquen su grado de eficiencia.

Asimismo, en los últimos años se ha debatido ampliamente si la medición debería centrarse en las diferencias en el estado de salud observadas entre grupos separados por factores socioeconómicos, geográficos o biológicos, o más bien en las inequidades, es decir, en aquellas desigualdades que son innecesarias, evitables, injustas e intolerables (2). La respuesta a esta pregunta, que debería buscarse en los objetivos de los diferentes estudios, tiene también implicaciones con respecto a cuáles son las medidas más adecuadas y al grado de agregación que deben tener los datos al calcularlas. La disponibilidad de datos desagregados y la necesidad de comparar diferentes

unidades geográficas también deben tenerse en cuenta a la hora de diseñar un estudio dirigido a medir desigualdades o diferencias entre grupos de población o entre individuos (3).

La idoneidad de las diferentes medidas para el estudio de las desigualdades e inequidades de salud ha sido examinada en diferentes estudios (4– 6). En el cuadro 1 se resumen las más comunes y se indica si incorporan el nivel socioeconómico, si pueden calcularse de forma individual o agregada, si son medidas absolutas o relativas y si son medidas de efecto (es decir, del efecto que el nivel socioeconómico tiene en la variable de

¹ Programa Especial de Análisis de Situación de Salud (SHA), Organización Panamericana de Salud, 525 23rd St. NW Washington DC, 20037 USA sha@paho.org

CUADRO 1. Características de las medidas de desigualdad utilizadas más frecuentemente

	Nivel socioeconómico	Agregado/ Individual	Medida de impacto	Relativa/ Absoluta
Cociente y diferencia				
Diferencia entre grupos	Sí	Ag/In	Sí/No	R
Cociente entre grupos	Sí	Ag/In	Sí/No	A
RAP ^a absoluto	Sí	Ag/In	Sí	R
RAP relativo	Sí	Ag/In	Sí	A
Derivadas de la curva de Lorenz				
Índice de Gini	No	In	–	R
Índice de Gini modificado	No	Ag/In	–	R
Índice de concentración	Sí	Ag/In	Sí	R
Índice de concentración generalizado	Sí	Ag/In	Sí	A
Basadas en regresión				
Regresión estándar				
Índice absoluto basado en regresión	Sí	Ag/In	No	R
Índice relativo basado en regresión	Sí	Ag/In	No	
Regresión de percentiles				
Índice de desigualdad de la pendiente	Sí	Ag/In	Sí	A
Índice relativo de desigualdad	Sí	Ag/In	Sí	R
RAP basado en regresión				
Versión absoluta	Sí	Ag/In	Sí	A
Versión relativa	Sí	Ag/In	Sí	R
Índice de disimilitud				
Índice de disimilitud absoluto	Sí	Ag/In	Sí	A
Índice de disimilitud relativo	Sí	Ag/In	Sí	R

Fuente: Referencia 7.

RAP: Riesgo atribuible poblacional; Ag: agregado; In: individual; A: absoluta; R: relativa.

salud [8]) o de impacto (si toman también en consideración el tamaño de los grupos afectados). El nivel de agregación de los datos y la incorporación o la ausencia de variables socioeconómicas en las mediciones de salud son dos elementos fundamentales sobre los cuales a menudo discrepan los profesionales que realizan estudios sobre la situación de salud de las poblaciones. Sus respectivas razones a favor y en contra se exponen a continuación.

Los defensores de medir las diferencias de salud (lo que supone la ausencia de variables socioeconómicas) en el menor nivel de agregación posible, es decir, entre individuos, subrayan que, en primer lugar, la información así obtenida puede complementar adecuadamente la información generada al calcular el valor promedio de un indicador de salud en una población; se entiende que el valor promedio de un indicador de salud, por su propia constitución, oculta parte de la información relevante para un estudio de desigualdades, la cual puede ser revelada mediante datos

individuales. En segundo lugar, la medición de diferencias individuales permite identificar a las personas en el peor estado de salud, independientemente de las variables utilizadas para definir distintos estratos o grupos socioeconómicos. En tercer lugar, la medición de diferencias entre individuos permite una comparación directa entre diferentes poblaciones siempre que los datos pertenezcan al mismo período, y también seguir la evolución de una misma población a lo largo del tiempo. Finalmente, evita los problemas asociados a la agregación de los datos (conformación de grupos, elección de variables socioeconómicas, cambios en la composición de los grupos a lo largo del tiempo). Las razones en contra de medir las diferencias de salud entre individuos son, en primer lugar, que el método pone de manifiesto el grado de desigualdad existente, pero no revela quiénes se ven afectados por causas socioeconómicas. También se aduce que el método está sujeto a la disponibilidad de datos con el máximo grado de desagregación y que en el caso de

variables dicotómicas, tales como la mortalidad, es necesario disponer de información en forma de variables continuas (tablas de vida de cohortes). Finalmente, se argumenta que si bien el método permite realizar mediciones de desigualdad entre individuos (9), estas no son fáciles de interpretar.

Los defensores de medir las desigualdades e inequidades de salud que se observan entre distintos grupos socioeconómicos argumentan que el estado de salud de una población está directamente relacionado con variables económicas y sociales, las cuales pueden ser tanto la causa como la consecuencia de dicho estado de salud. Una razón para utilizar datos agrupados es que el resultado de la medición de desigualdades e inequidades que con ello se obtiene no solo indica la magnitud de la desigualdad, sino también en qué grupos se observa (dónde se sitúa geográfica, biológica o socialmente), información que debería utilizarse para llevar a cabo acciones públicas más eficaces, tanto en la elección de los problemas cuya solución es más urgente, como en la focalización adecuada de las acciones encaminadas a afrontarlos. Con este enfoque es posible hacer comparaciones dentro de un mismo grupo poblacional a lo largo del tiempo si no se cambian los parámetros de construcción de los grupos socioeconómicos, o entre diferentes poblaciones si se utilizan grupos reducidos de un tamaño similar que garantice su homogeneidad interna, aunque esta metodología aún no tiene el respaldo de suficientes pruebas empíricas acumuladas. Asimismo, son varios los argumentos esgrimidos en contra de medir las desigualdades e inequidades entre diferentes grupos socioeconómicos: la pérdida de información que encierran los valores medios resultantes de la agregación de los individuos en grupos; la dificultad de comparar diferentes poblaciones debido a la heterogeneidad de las variables elegidas para la construcción de grupos socioeconómicos; la variación a lo largo del tiempo en la composición de dichos grupos, y el hecho de que la

construcción de grupos socioeconómicos podría eclipsar la importancia de otros posibles factores determinantes de las desigualdades de salud en el nivel individual.

Este artículo ilustra, mediante un ejemplo, las diferencias entre los resultados que arrojan las medidas de desigualdad utilizadas con mayor frecuencia cuando se usan dos tipos de agregación: socioeconómica y geográfica. En cada caso se construyen partiendo de dos niveles distintos de agregación (provincias y cantones).

MATERIALES Y MÉTODOS

El ejemplo que se presenta para ilustrar el estudio de desigualdades hace referencia a las tasas de mortalidad infantil (TMI) de Costa Rica de 1973 a 1984. Cuatro son las agregaciones para las cuales se calcula el nivel de desigualdad de dichas tasas: 1) provincias (nivel geográfico) 2) cantones (nivel geográfico) 3) estratos socioeconómicos construidos a partir de datos provinciales 4) estratos socioeconómicos construidos a partir de datos cantonales.

Mientras que entre 1973 y 1984 el número de provincias de Costa Rica permaneció invariable (siete), el número de cantones ascendió de 79 en 1973 a 81 en 1984 (Garabito y Corredores se escindieron durante ese período, siendo agrupados en el estudio en sus cantones de origen para facilitar las comparaciones). Los 79 cantones se han reducido en el estudio finalmente a 62 (1973) y a 64 (1984), habiéndose aplicado el criterio de que las unidades básicas debían tener un mínimo de 300 nacimientos por año con el objetivo de evitar distorsiones al trabajar con números pequeños, obstáculo que tiene especial relevancia en los estudios de desigualdades. Los datos de los cantones que no cumplían dicho criterio fueron agrupados en otros cantones bajo el mismo nivel socioeconómico y siguiendo criterios de proximidad geográfica.

La construcción de los niveles socioeconómicos se llevó a cabo utilizando las variables de salario y nivel de

CUADRO 2. Tasas de mortalidad infantil máximas y mínimas (por 1 000 nacidos vivos) para los diferentes niveles de agregación. Costa Rica, 1973 y 1984

	Agregación por nivel socioeconómico		Agregación por nivel geográfico	
	1973	1984	1973	1984
Unidad geográfica básica: provincia				
Número de grupos	4	4	7	7
Tasa máxima	55,34	20,02	65,29	27,47
Nacimientos (No.)	13 932	27 769	4 166	6 370
Defunciones (No.)	771	556	272	175
Tasa mínima	33,80	16,72	33,64	15,97
Nacimientos (No.)	18 847	27 327	18 847	13 279
Defunciones (No.)	637	457	637	212
Unidad geográfica básica: cantón				
Número de grupos	4	4	62	64
Tasa máxima	51,67	19,31	87,18	44,86
Nacimientos (No.)	11 110	29 162	585	535
Defunciones (No.)	574	563	51	24
Tasa mínima	33,24	17,05	17,26	2,92
Nacimientos (No.)	10 560	12 493	753	342
Defunciones (No.)	351	213	13	1

educación, igualmente ponderadas con datos procedentes del Censo de Población y Vivienda de Costa Rica (1973) y ordenadas aplicando un índice de 0 a 1, metodología basada en la propuesta por el Programa de las Naciones Unidas para el Desarrollo (PNUD) para construir el índice de desarrollo humano. Dicha metodología, inspirada en los trabajos de S. Anand y A. Sen y aplicada en el *Informe de desarrollo humano* de 1999 (11), implica utilizar el logaritmo del ingreso en lugar de su valor absoluto. A partir de dicha ordenación se tomaron cuartiles para dividir la población en cuatro grupos socioeconómicos.

El número de nacimientos y defunciones en cada una de las cuatro agrupaciones estudiadas, así como sus respectivas tasas máximas y mínimas en las unidades básicas de agrupación, se presentan en el cuadro 2.

Se calcularon valores para 17 medidas de desigualdad que se clasificaron en tres grupos: 1) las basadas en cocientes y diferencias entre grupos;

2) las derivadas de la curva de Lorenz junto con los índices de disimilitud, y 3) las basadas en modelos de regresión. Para cada una de las medidas de desigualdad se calcularon 8 valores (4 para 1973 y 4 para 1984) a partir de los

cuatro ejemplos de agregación de los datos: desigualdades entre los cuatro niveles socioeconómicos basados en datos de provincias; desigualdades entre los cuatro niveles socioeconómicos basados en datos de cantones; desigualdades entre las 7 provincias, y desigualdades entre los 62 (1973) ó 64 (1984) cantones de Costa Rica.

RESULTADOS

Medidas basadas en cocientes y diferencias

Las medidas basadas en cocientes y diferencias (cuadro 3) se utilizan muy frecuentemente en los estudios de desigualdades por ser fáciles de calcular e interpretar. Sin embargo, al utilizar cocientes y diferencias entre grupos socioeconómicos extremos o usando tasas máximas y mínimas, se omite la información que aporta el resto de los grupos. Como muestra el ejemplo, es habitual que los datos agregados por niveles socioeconómicos revelen grados de desigualdad menores que los agregados por unidades geográficas, debido a que estas constituyen un número mayor de grupos o unidades básicas. Una menor agregación encierra

CUADRO 3. Resultados de las medidas de desigualdad basadas en cocientes y diferencias. Costa Rica, 1973 y 1984

	Agregación por nivel socioeconómico			Agregación por nivel geográfico		
	1973	1984	Diferencia en el índice de desigualdad (1973–1984)	1973	1984	Diferencia en el índice de desigualdad (1973–1984)
Número de grupos						
Unidad básica: provincia	4	4		7	7	
Unidad básica: cantón	4	4		62	64	
Diferencia entre grupos socioeconómicos extremos						
Unidad básica: provincia	21,54	2,28	19,26	22,03	3,32	18,71
Unidad básica: cantón	18,43	-0,73	19,16	16,93	3,77	13,16
Cociente entre grupos socioeconómicos extremos						
Unidad básica: provincia	1,64	1,14	0,50	1,65	1,20	0,46
Unidad básica: cantón	1,55	0,96	0,60	1,54	1,29	0,25
Diferencia entre grupos con tasas máxima y mínima						
Unidad básica: provincia	21,54	3,30	18,24	31,65	11,51	20,14
Unidad básica: cantón	18,43	2,26	16,17	69,92	41,94	27,98
Cociente entre grupos con tasas máxima y mínima						
Unidad básica: provincia	1,64	1,20	0,44	1,94	1,72	0,22
Unidad básica: cantón	1,55	1,13	0,42	5,05	15,34	-10,29
RAP ^a absoluto						
Unidad básica: provincia	10,38	1,75	8,63	10,54	1,75	8,79
Unidad básica: cantón	10,94	1,42	9,52	12,85	1,71	11,14
RAP relativo (%)						
Unidad básica: provincia	23,50	9,46	14,03	23,86	9,46	14,39
Unidad básica de datos: cantón	24,77	3,73	21,03	29,10	9,27	19,83

^a RAP: Riesgo atribuible poblacional.

la posibilidad de comparar grupos relativamente pequeños con tasas notablemente altas o bajas, que de otra forma quedarían diluidas en los valores medios de unidades geográficas de superior nivel. No es de sorprender, por tanto, que incluso en situaciones generales de cierta igualdad, como en el caso de la TMI de Costa Rica en 1984, el cociente entre las tasas máxima y mínima de los grupos geográficos divididos por cantones arroje un valor de 15,34, que triplica el obtenido para 1973, debido a que un cantón tuvo una TMI muy baja en relación con el resto.

Problemas como el subregistro, que es una de las posibles causas de que haya cifras demasiado bajas, quedan ocultos cuando se utilizan datos agregados en los entornos donde se produce. Situaciones similares se presentan en el caso del riesgo atribuible poblacional (RAP). Los resultados de 1984 indican que la

agregación por grupos socioeconómicos basada en cantones mostraba una desigualdad menor, tanto en términos absolutos como relativos (1,42 y 3,73%, respectivamente) que si se hubiera utilizado la desagregación por unidad geográfica, es decir, por cantones (1,71 y 9,27%, respectivamente).

Debe tenerse en cuenta que si se elige comparar la población con el grupo más aventajado socialmente (y no con el que goza de un mejor estado de salud), una mayor desagregación conlleva una menor probabilidad de que el grupo más aventajado socialmente sea también el que está en mejor estado de salud. Esto implica que la diferencia entre la media y el grupo de referencia puede generar información que, aun siendo verdadera, no corresponde a la máxima reducción posible del

indicador de salud y, por lo tanto, no muestra la máxima desigualdad existente entre los grupos poblacionales.

Medidas derivadas de la curva de Lorenz y el índice de disimilitud

El índice de Gini modificado² utiliza datos agrupados en vez de individuales (como ocurre con el índice de Gini) y ordena los grupos según una variable que representa su estado de salud. Por lo tanto, dicho índice es particularmente sensible a la agregación de los datos y su valor es igual o mayor cuanto mayor sea la desagregación (cuadro 4). Como se observa en la figura 1, del grado de agregación depende que se obtengan resultados muy distintos: la agregación

² En inglés este índice, que es una adaptación del índice de Gini, se conoce por *pseudo-Gini index*.

³ Esta curva, que es una adaptación de la de Lorenz, se conoce en inglés por *pseudo-Lorenz curve*.

CUADRO 4. Resultados de medidas de desigualdad derivadas de la curva de Lorenz y de los índices de disimilitud

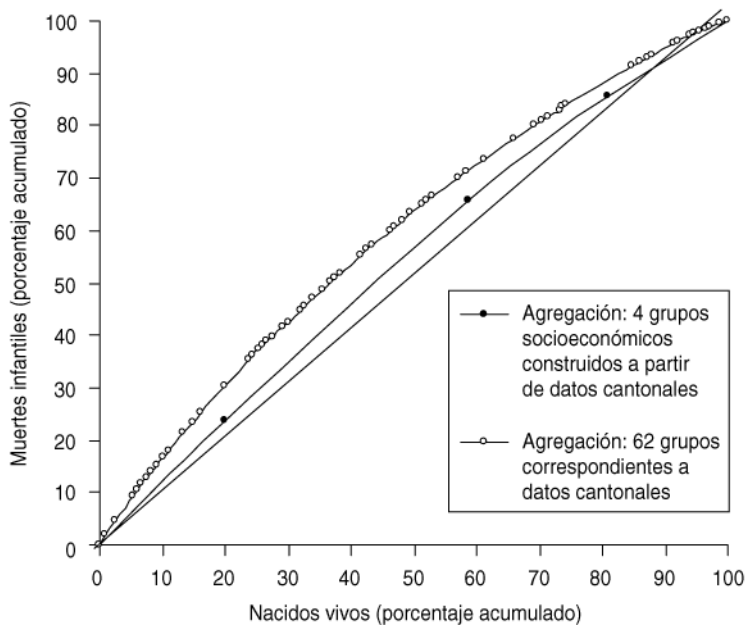
	Agregación por nivel socioeconómico			Agregación por nivel geográfico		
	1973	1984	Diferencia en el índice de desigualdad (1973–1984)	1973	1984	Diferencia en el índice de desigualdad (1973–1984)
Número de grupos						
Unidad básica: provincia	4	4		7	7	
Unidad básica: cantón	4	4		62	64	
Índice de Gini modificado						
Unidad básica: provincia	0,11	0,04	0,07	0,13	0,08	0,05
Unidad básica: cantón	0,08	0,02	0,06	0,19	0,14	0,05
Índice de concentración						
Unidad básica: provincia	0,11	0,03	0,08	0,10	0,02 ^a	0,08
Unidad básica: cantón	0,08	0,006	0,07	0,10	0,04 ^a	0,06
Índice de concentración generalizado						
Unidad básica: provincia	4,86	0,55	4,31	4,42	0,37 ^a	4,05
Unidad básica: cantón	3,53	0,11	3,42	4,42	0,74 ^a	3,68
Índice de disimilitud absoluto						
Unidad básica: provincia	216	51	165	285	91	194
Unidad básica: cantón	169	28	141	336	134	202
Índice de disimilitud relativo (%)						
Unidad básica: provincia	8,88	3,65	5,23	11,71	6,46	5,24
Unidad básica: cantón	6,93	2,00	4,93	13,80	9,57	4,23

^a Esta curva de concentración cruza la diagonal.

de los datos cantonales en cuatro grupos socioeconómicos para 1973 arroja un índice de Gini modificado de 0,08, el cual asciende hasta 0,19 cuando la información no se agrega y se utilizan los datos individuales de los 62 cantones para construir la curva de Lorenz modificada.³

A diferencia de los índices de Gini y de Gini modificado, para el índice de concentración los grupos se ordenan por nivel socioeconómico, con un rango de [-1,1] (Gini, Gini modificado [0,1]). Por ello hay que tener cautela al interpretar valores cercanos a cero, que pueden indicar la ausencia de un gradiente por niveles socioeconómicos (y no la ausencia de desigualdades), o ser el resultado de cruces de las curvas con la diagonal, haciendo que las áreas a uno y otro lado de esta se compen-sen (cuadro 4), como ocurre en el caso de la agregación geográfica de los datos de 1984, tanto por provincias como por cantones (0,02 y 0,04, respectivamente). Mientras más desagregados estén los datos, mayor es la probabilidad de que se produzcan alteraciones en su ordenamiento por estado de salud y, en

FIGURA 1. Índices de Gini modificados. Tasas de mortalidad infantil (por 1 000 nacidos vivos) basadas en datos cantonales. Costa Rica, 1973



menor medida, de que se produzcan cruces de la diagonal que dificultan la interpretación y comparación del

índice. Por otra parte, el índice no necesariamente aumenta cuando es mayor la desagregación de los datos de

CUADRO 5. Resultados de medidas de desigualdad basadas en modelos de regresión

	Agregación por nivel socioeconómico			Agregación por nivel geográfico		
	1973	1984	Diferencia en el índice de desigualdad (1973–1984)	1973	1984	Diferencia en el índice de desigualdad (1973–1984)
Número de grupos						
Unidad básica: provincia	4	4		7	7	
Unidad básica: cantón	4	4		62	64	
Regresión de percentiles						
Índice de desigualdad de la pendiente						
Unidad básica: provincia	-31,26	-4,21	-27,05	-26,77	-2,59	-24,18
Unidad básica: cantón	-23,63	0,07	-23,70	-25,36 ^a	0,44 ^a	-25,80
Regresión de percentiles						
Índice relativo de desigualdad						
Unidad básica: provincia	2,09	1,26	0,84	1,87	1,15	0,72
Unidad básica: cantón	1,73	1,00	0,73	1,81 ^a	0,98 ^a	0,83
RAP ^b absoluto basado en regresión (TMI ^c frente a años de educación) para el grupo socioeconómico más aventajado						
Unidad básica: provincia	10,04	1,29	8,75	10,61	2,23	8,38
Unidad básica: cantón	11,67	0,17	11,50	24,24 ^a	3,21 ^a	21,03
RAP relativo (%) basado en regresión (TMI frente a años de educación) para el grupo socioeconómico más aventajado						
Unidad básica: provincia	22,73	6,99	15,73	24,02	12,08	11,93
Unidad básica: cantón	26,41	0,93	25,49	54,87 ^a	17,39 ^a	37,48
Regresión estándar						
Índice absoluto (TMI frente a años de educación)						
Unidad básica: provincia	-12,46	-1,46	-11,00	-14,13	-2,98	-11,15
Unidad básica: cantón	-6,48	0,06	-6,54	-9,09 ^a	-0,93 ^a	-8,16
Regresión estándar						
Índice relativo (TMI frente a años de educación) para el grupo socioeconómico más aventajado						
Unidad básica: provincia	0,23	0,08	0,15	0,25	0,15	0,11
Unidad básica: cantón	0,13	0,00	0,12	0,19 ^a	0,07 ^a	0,12

^a Ajuste de la recta de regresión sin significación estadística.

^b RAP: riesgo atribuible poblacional.

^c TMI: tasa de mortalidad infantil.

origen. En lo que respecta al índice de disimilitud, los resultados del ejemplo indican que de 216 (8,88%) a 336 (13,80%) defunciones —estas cifras corresponden a la agregación socioeconómica por provincias y la agregación geográfica por cantones, respectivamente— deberían ser redistribuidas para lograr que los diferentes grupos dentro de cada nivel de agregación tuvieran la misma TMI para 1973.

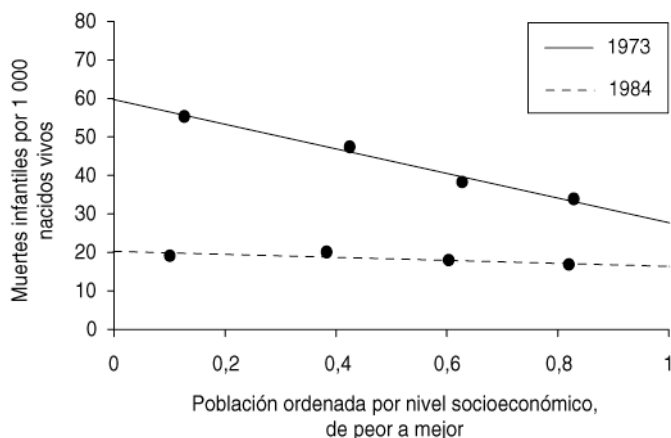
Medidas basadas en modelos de regresión

Los tres tipos de medidas basadas en modelos de regresión que ilustra el ejemplo (cuadro 5) —regresión de percentiles (índice de desigualdad de la pendiente e índice relativo de desigualdad), el RAP basado en una regresión, y la regresión estándar (estos dos últimos en sus versiones absoluta y relativa)— tienen como principal dificultad que deben cumplirse las hipótesis aplicables a los modelos de regresión, aunque en estos casos tengan

una finalidad descriptiva en lugar de predictiva. Asimismo, en el caso de series de datos con una marcada variabilidad, mientras mayor sea la desagregación, menor será el ajuste de la recta de regresión y los resultados podrían considerarse estadísticamente no significativos, como ocurre en las agregaciones por niveles geográficos cantonales, tanto en 1973 como en 1984 (cuadro 5).

En la regresión estándar, cuando los grupos socioeconómicos se basan en datos provinciales, el índice absoluto (resultado de la regresión de la TMI

FIGURA 2. Regresión de percentiles para la tasa de mortalidad infantil entre grupos socioeconómicos a partir de provincias. Costa Rica, 1973, 1984



frente a los años de educación) indica que en 1973 cada año adicional de educación implicaba una disminución anual de 12,46 muertes infantiles (es decir, de menores de un año). En cambio, cuando los grupos socioeconómicos se construyen a partir de datos cantonales, el índice absoluto indica que cada año adicional de educación implicaría una reducción de solamente 6,48 muertes infantiles en el mismo período. Cuando se usa el RAP basado en una regresión, una mayor desagregación aumenta la probabilidad de que la diferencia entre el RAP observado y el RAP basado en la regresión sea mayor (ya que es más alta la probabilidad de que el dato de referencia se encuentre alejado de la recta de regresión). En el ejemplo, el RAP que es resultado de la agregación socioeconómica basada en cantones para 1973 es 26,41%, mientras que el valor correspondiente al RAP que no se basa en una regresión es de 24,77% para ese mismo año. También al aplicarse una regresión de percentiles se aprecian diferencias según la agregación de los datos, pero en este caso en sentido opuesto: cuando se miden las desigualdades entre los cuatro grupos socioeconómicos definidos a partir de datos provinciales, la diferencia entre la base y la cima de la jerarquía socioeconómica es de 31,26 muertes anuales de menores de un año (figura

2), mientras que la misma medida basada en datos cantonales arroja una diferencia de 23,63 muertes de menores de un año. Una hipótesis que podría explicar esta inversión podría ser la falta de una clara correlación entre el nivel socioeconómico y la TMI en el nivel cantonal; otra posibilidad es la presencia de un subregistro de la mortalidad en determinadas zonas.

Cuando las tasas de mortalidad y morbilidad son relativamente parecidas y pequeñas, no es de sorprender que las medidas de desigualdad relativas generen valores altos causados por grandes diferencias relativas con respecto a la medida absoluta. Así por ejemplo, aun cuando la diferencia absoluta es de solo 3,32 muertes infantiles por 1 000 nacidos vivos en la agregación por grupos socioeconómicos (1984), el cociente de las tasas de mortalidad infantil entre grupos socioeconómicos extremos indica que el grupo socioeconómico más aventajado tiene una mortalidad infantil 20% menor. Asimismo, no debe sorprender que los valores obtenidos mediante las medidas de desigualdad entre niveles socioeconómicos creados a partir de la comparación entre provincias y cantones sean en todos los casos distintos, ya que las poblaciones pertenecientes a cada nivel socioeconómico son diferentes en ambos casos: mientras que en el

primero todos los cantones de una provincia son tratados como si perteneciesen a un mismo nivel socioeconómico, en el segundo cada cantón es asignado al nivel socioeconómico que le corresponde.

Los resultados obtenidos indican que las medidas de desigualdad presentadas (a excepción de las que se basan en modelos de regresión) no parecen ser sensibles al grado de agregación de los datos utilizados para formar grupos socioeconómicos. Asimismo, los valores de las medidas de desigualdad entre los grupos agregados por unidades geográficas tienden a ser mayores cuanto mayor sea la desagregación, siempre que no haya ninguna ordenación por niveles socioeconómicos.

DISCUSIÓN

Antes de realizar un estudio de desigualdades e inequidades en materia de salud, se deben tomar algunas decisiones relativas al diseño del estudio que, como muestra el ejemplo, pueden ejercer gran influencia sobre sus resultados. Entre ellas merecen especial énfasis la elección de las medidas de desigualdad y la decisión del grado de agregación de los datos, siempre en función de los objetivos del estudio y de los datos disponibles. En caso de llevarse a cabo una agregación socioeconómica, es preciso seleccionar las variables y la metodología que se usarán para crear los grupos socioeconómicos. No debe olvidarse que si bien las diferentes medidas y agregaciones generan diferentes tipos de información que se complementan entre sí, no todos se adecuan a los objetivos de los distintos estudios.

En los estudios cuyo objetivo no solo es describir las desigualdades sociales en el ámbito de la salud, sino también servir de herramienta para emprender acciones posteriores, se deberían utilizar medidas que contengan variables socioeconómicas y hacer una agregación de los datos. Aunque las mediciones relacionadas con grupos adolecen de los problemas que genera la agregación —variaciones en el

tamaño de los grupos a lo largo del tiempo, definiciones distintas para la creación de los grupos socioeconómicos—, permiten establecer hipótesis acerca de las posibles causas y consecuencias de las desigualdades e inequidades en materia de salud, ya que abarcan factores sociales, geográficos y comunitarios de tipo contextual, lo cual puede orientar soluciones de salud pública a problemas que afectan a un determinado número de personas.

Al decidir qué grado de agregación de los datos es el más adecuado, se deben sopesar las consecuencias de una mayor o menor agregación. La agregación comporta habitualmente una pérdida de información debido a la dilución de las medidas resumen⁴ de cada grupo que se agrega (habitualmente la media), de modo que puede ocultar (en caso de que existan) tasas extremas que afecten a una población relativamente pequeña dentro de la unidad básica de agregación escogida. Una mayor agregación supone, asimismo, una mayor coherencia de los resultados obtenidos, aunque es importante subrayar que el resultado del análisis de desigualdades e inequidades corresponderá a las diferencias existentes entre las unidades de agregación, por lo que reviste una importancia fundamental indicar no solo cuáles son las unidades básicas, sino qué grado de agregación se ha aplicado para obtenerlas. Los análisis con unidades básicas más desagregadas (es decir, menores) se efectúan con el fin de conocer mejor las diferencias entre grupos más pequeños de población y, por tanto, de entender más exhaustivamente los problemas de salud en poblaciones pequeñas, lo cual redundará en una mayor homogeneidad relativa dentro de cada una de las unidades. Sin embargo, es preciso recordar las dificultades estadísticas que entraña trabajar con números pequeños en la medición de desigualdades, así como la posible pérdida de significación estadística al

buscar medidas sintéticas para series relativamente grandes de valores muy dispares, como se ha observado al efectuar las regresiones con datos cantonales en el ejemplo, donde los modelos tenían un poder explicativo limitado.

Los estudios destinados a conocer las diferencias de salud entre individuos o a describir situaciones de salud sin que necesariamente se proyecte ninguna acción posterior derivada del estudio pueden basarse en datos individuales sin asociaciones con variables de tipo socioeconómico. Las mediciones de las variaciones observadas en la salud individual (por ejemplo, diferencias en la longevidad de las personas mediante tablas de vida) constituyen una información en extremo interesante para buscar una medida global de desigualdad que no encierre los problemas relacionados con las construcciones socioeconómicas. No obstante, dichas mediciones no parecen estar encaminadas —al menos no de forma directa— en la dirección de un análisis que sirva de base para realizar acciones concretas de salud pública.

Con relación a las medidas propiamente dichas, es preferible utilizar el cociente y la diferencia entre las tasas solamente cuando los datos están agregados, ya que aportan poca información sobre la situación global cuando las unidades básicas son pequeñas. En medidas de RAP, debe tomarse en cuenta que si se utiliza como referente únicamente el grupo socioeconómico más favorecido, mientras mayor sea la agregación, mayor será la probabilidad de que el resultado no refleje la máxima reducción posible de la variable, por lo que la desigualdad obtenida será menor que la verdadera en el nivel de agregación escogido. El índice de Gini modificado es especialmente sensible a la agregación de datos y mostrará una desigualdad igual o mayor mientras más desagregados estén los grupos, pero ello no siempre se produce en el caso del índice de concentración y del índice de concentración generalizado, dada su ordenación por niveles socioeconómicos. Cuando los índices se

basan en modelos de regresión se recomienda utilizar datos agregados con el objetivo de que dichos modelos expliquen de forma robusta las series de datos, cumplan las restricciones de los modelos y fortalezcan los resultados.

Para que el conocimiento de la situación de salud resulte más completo, es imprescindible escoger diversas medidas que aporten visiones diferentes y complementarias sobre la misma situación de salud. Con ello no solo se enriquecerá el análisis, sino que también se perfilará mejor el tipo de acciones que se han de llevar a cabo, si este es uno de los objetivos del estudio. Asimismo, es importante tener en cuenta quiénes son los destinatarios de la información —si son políticos o técnicos— a fin de adaptar de forma consecuente la información generada en función de la complejidad de su interpretación, procurando así lograr que el proceso sea más eficaz.

Todavía queda un largo camino que recorrer en lo que concierne a la valoración de las desigualdades que se pueden considerar inequidades. Aunque ha habido algunas aproximaciones iniciales en la medición de diferencias individuales, falta que estas se extiendan a la medición de diferencias entre grupos, tomando como base algunas de las mediciones efectuadas por los economistas para cuantificar la desigualdad de los ingresos. Por ejemplo, es necesario determinar si debe valorarse la salud de los ricos de la misma manera que la de los pobres (y en caso negativo, cuál ha de ser la relación), o cuándo y en qué grado la mejora del nivel de salud general debe sacrificarse para obtener la reducción del grado de desigualdad en una situación de salud. Ambos conceptos —y no solo el binomio igualdad-desigualdad— forman parte de las percepciones de las personas acerca de las situaciones de salud, lo cual debe permitirnos determinar qué situaciones pueden considerarse poco equitativas y, por tanto, constituir un objetivo prioritario de las acciones y cooperaciones técnicas de las organizaciones públicas en el ámbito de la salud.

⁴ Estas medidas también se conocen por “medidas sintéticas” (*summary measures* en inglés).

REFERENCIAS

1. Organización Panamericana de la Salud. En busca de la equidad. Informe anual del Director, 1995. Washington, D.C.: OPS; 1996. (Documento Oficial No. 277).
2. Whitehead M. The concepts and principles of equity and health. Copenhagen: World Health Organization, Regional Office for Europe; 1990.
3. Murray CJL, Gakidou, Frenk J. Health inequalities and social group differences: what should we measure? *Bull World Health Organ* 1999;77(7):537-543.
4. Wagstaff A, Paci P, Doorslaer E. On the measurement of inequalities in health. *Soc Sci Med* 1991;33:545-557.
5. Kunst AE, Mackenbach JP. Measuring socioeconomic inequalities in health. Copenhagen: World Health Organization, Regional Office for Europe; 1994.
6. Pamuk ER, Lenzner H, Brackbill R. Measuring socioeconomic inequality in health: an update on methodological issues. Proceedings of the 1993 Public Health Conference on Records and Statistics. Washington, D.C.: National Center for Health Statistics; 1993.
7. Borrell C, Rué M, Pasarín MI, Benach J, Kunst JAE. La medición de las desigualdades en salud. *Gaceta Sanitaria* 2000;14(Supl. 3):20-33.
8. Kunst A, Mackenbach JP. Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. *Soc Sci Med* 1997;44:757-771.
9. Gakidou EE, Murray CJL, Frenk J. Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bull World Health Organ* 2000;78:42-54.
10. Braveman P, Krieger N, Lynch J. Health inequalities and social inequalities in health. *Bull World Health Organ* 2000;78:232-234.
11. United Nations Development Programme. Human Development Report 1999. New York: UN; 2000.

Manuscrito recibido el 8 de febrero de 2002. Aceptado para publicación, tras revisión, el 15 de agosto de 2002.

ABSTRACT

Data aggregation in measuring inequalities and inequities in the health of populations

Objectives. To compare how different degrees of data aggregation influence the measurement of health inequalities and health inequities within a population, and to assess the appropriateness of those different degrees of data aggregation in performing studies on inequalities and inequities.

Methods. As an example, we used data on the infant mortality rate in Costa Rica in 1973 and in 1984 and calculated measurements that are frequently used to quantify inequalities and inequities.

Results. According to our results, the inequality measures presented (except for those that were derived using regression models) are not sensitive to data aggregation by socioeconomic groups. However, when geographic areas are compared, more disaggregation of the data results in the measures indicating greater inequality.

Conclusions. Our results show that some measures can vary widely depending on the level of data aggregation. It is thus crucial to know how to select these measures and also how to aggregate the data in a way that is consistent with the objectives of each study.
