

B1: PROBLEMAS

1) Invente unos datos con 3 variables (Y, respuesta; X, intervención; Z, condición) dicotómicas en los que Z sea confusora.

2) Invente unos datos con 2 variables dicotómicas (X, intervención; Z, condición) y una respuesta Y numérica en los que Z sea confusora.

3) Invente una situación hipotética con 3 variables (Y, respuesta; X, intervención; Z, condición) en la que Z pueda ser una confusora verosímil. Dibuje el DAG.

4) PELIRROJAS: ¿Tienen más miedo las pelirrojas en el dentista? Lea los siguientes documentos y:

- Interprete la tabla de resultados.
- Comente sobre la variable principal.
- ¿Se trata de un estudio de predicción o de intervención?
- ¿Qué variable se asigna al azar?
- La conclusión, ¿tiene el valor de un estudio observacional o experimental?

La nota de agencia de Europa Press del 26/12/12 titulada *"Las mujeres pelirrojas sufren más miedo a acudir al dentista"* incluye en el texto:

Las mujeres pelirrojas sufren más miedo a acudir al dentista que las morenas, tal y como ha evidenciado una investigación publicada en la edición de diciembre de la revista 'Journal of Endodontics' y que recoge la plataforma 'miedoaldolor.com'. Así, este estudio, ha comprobado tras analizar a 124 mujeres, morenas y pelirrojas en la misma proporción, que las segundas sufren más miedo al acudir a este profesional sanitario que las otras. El color rojizo del cabello "está relacionado con una variación del gen receptor de la melanocortina-1 (MC1R)", lo que podría ser la explicación, aseguran los expertos. En este sentido, indican que el pelo rojo y el gen MC1R "estuvieron significativamente vinculados a mayores niveles de ansiedad dental". Además, trabajos anteriores "han demostrado que la presencia de la variación genética del MC1R estaba relacionada con una mayor sensibilidad al dolor y la analgesia", señalan.

Por su lado, el *Journal of Endodontics* (J Endod 2012;38:1564–1569) ha publicado el siguiente resumen y tabla de resultados:

Introduction: The exact reasons for failure of the inferior alveolar nerve (IAN) block are not completely known, but red hair could play a role. The genetic basis for red hair involves specific mutations, red hair color (RHC) alleles, in the melanocortin-1 receptor (MC1R) gene.

The purpose of this prospective randomized study was to investigate a possible link between certain variant alleles of the MC1R gene or its phenotypic expression of red hair and the anesthetic efficacy of the IAN block in women.

Materials: One-hundred twenty-four adult female subjects (62 red haired and 62 dark haired) participated in this study. Dental anxiety was determined in each subject using the Corah Dental Anxiety Questionnaire. The subjects were given 2 cartridges of 2% lidocaine with 1:100,000 epinephrine via the IAN block. Pulpal anesthesia was measured in the posterior and anterior teeth in 4-minute cycles for 60 minutes using an electric pulp tester. The MC1R alleles were genotyped for each subject from cheek cells containing DNA collected using buccal swabs. *Results:* Women with red hair and women with 2 RHC alleles reported significantly higher levels of dental anxiety compared with women with dark hair or women with 0 RHC alleles. No significant differences in anesthetic success were found between any of the groups for any of the teeth. *Conclusions:* Red hair and the MC1R gene were significantly linked to higher levels of dental anxiety but were unrelated to success rates of the IAN block in women with healthy pulps.

TABLE 4. Discomfort Ratings (VAS) of IAN Block

Group (mean \pm SD mm)	None, % (n)	Mild, % (n)	Moderate, % (n)	Severe, % (n)	P value
Needle insertion					
Red-haired group (n = 62)					
59.3 \pm 26.3	0 (0)	48 (30)	52 (32)	0 (0)	
Dark-haired group (n = 62)					
44.2 \pm 24.9	5 (3)	66 (41)	29 (18)	0 (0)	.048*
2 RHC alleles (n = 45)					
61.3 \pm 27.0	0 (0)	47 (21)	53 (24)	0 (0)	
0 RHC alleles (n = 52)					
45.7 \pm 25.1	4 (2)	65 (34)	31 (16)	0 (0)	.090
Needle placement					
Red-haired group (n = 62)					
60.3 \pm 31.0	3 (2)	44 (27)	50 (31)	3 (2)	
Dark-haired group (n = 62)					
52.4 \pm 28.3	5 (3)	53 (33)	42 (26)	0 (0)	.668
2 RHC alleles (n = 45)					
62.6 \pm 29.1	2 (1)	42 (19)	53 (24)	2 (1)	
0 RHC alleles (n = 52)					
51.0 \pm 28.2	4 (2)	58 (30)	39 (20)	0 (0)	.368
Solution deposition					
Red-haired group (n = 62)					
72.9 \pm 34.6	0 (0)	37 (23)	48 (30)	15 (9)	
Dark-haired group (n = 62)					
63.7 \pm 30.9	0 (0)	47 (29)	50 (31)	3 (2)	.498
2 RHC alleles (n = 45)					
72.1 \pm 34.6	0 (0)	36 (16)	51 (23)	13 (6)	
0 RHC alleles (n = 52)					
62.8 \pm 31.7	0 (0)	50 (26)	44 (23)	6 (3)	.613

SD, standard deviation.

*There was a significant difference ($P < .05$) between the groups.

En el texto, los únicos apartados que mencionan la aleatorización son:

Before the experiment, the red-haired and dark-haired female subjects were randomly assigned 6-digit numbers from a random number table. Each subject was also randomly assigned a right or left designation to determine which side of the mandible would receive the IAN block. Only the random numbers were recorded on the data collection sheets and genotype samples to further blind the experiment.

The tubes were labeled with the random 6-digit numbers to keep the subjects' identity and hair trait confidential from the laboratory. The tubes were packed and shipped to the DNA Core at the Center for Genetics and Molecular Medicine at the University of Louisville, Louisville, KY, for processing. This laboratory was chosen because they had experience genotyping the MC1R gene (8, 9). Only the MC1R genes were analyzed and genotyped.

Between-group comparisons for the red-haired subjects and dark-haired subjects for age and side of the mandible were performed using the randomization test and chi-square test, respectively.

Finalmente, sobre la variable respuesta, la escala visual analógica (EVA), dice:

Subjects were instructed before the study on how to rate the pain for each phase of the first IAN block including needle insertion, needle placement, and deposition of anesthetic solution using a Heft-Parker visual analog scale (VAS). The VAS was divided into 4 categories. No pain corresponded to 0 mm. Mild pain was defined as greater than 0 mm and less than or equal to 54 mm. Mild pain included the descriptors of faint, weak, and mild pain. Moderate pain was defined as greater than 54 mm and less than 114 mm. Severe pain was defined as equal to or greater than 114 mm. Severe pain included the descriptors of strong, intense, and maximum possible

5) ORDENACIÓN DE LOS DISEÑOS CIENTÍFICOS. La academia de ciencias médicas escandinavas propone la siguiente ordenación de los diseños habituales de acuerdo con la calidad o solidez de la evidencia que aportan:

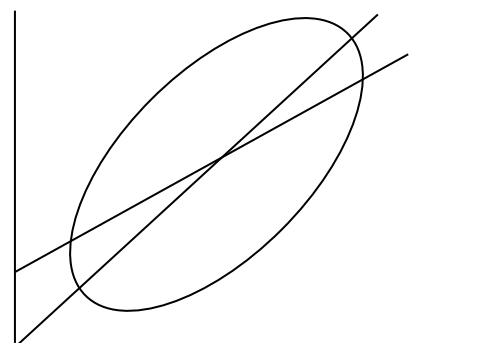
- a) Ensayo controlado y aleatorizado con muestra grande
- b) Ensayo controlado y aleatorizado con muestra pequeña
- c) Ensayo no aleatorizado con controles coincidiendo en el tiempo
- d) Ensayo no aleatorizado con controles históricos
- e) Estudio prospectivo ("de cohortes")
- f) Estudio retrospectivo ("de casos-control")
- g) Estudio transversal
- h) Estudios sobre bases de datos o registros de vigilancia epidemiológica
- i) Serie consecutiva de casos.
- j) Notificación de un caso aislado ("anécdota").

Busque información sobre qué significa cada uno y justifique el orden propuesto. Hay mucha información sobre estos estudios en la red. Se trata de escoger y resumir algún argumento que permita decir porqué cada uno es mejor que el siguiente (la razón será diferente en muchos casos). Por ejemplo, "a>b ya que tendrá un error típico menor y la información sobre la magnitud del efecto de la intervención será más precisa (IC95% más estrecho)".

6) GANANCIAS DE EMPRESAS (Wonnacott).

Busque por internet información sobre el fenómeno de regresión a la media (p.e., <http://bioestadistica.upc.edu/MaterialDocente/Modulo4>; capítulo 10, punto 6) y responda el siguiente ejercicio.

Suponga que todas las empresas de cierto sector industrial tuvieron las ganancias de 1999 y 2000 como indica el gráfico, donde la recta de regresión por mínimos cuadrados es $G_{2000} = 0.04 + 0.5 \cdot G_{1999}$.



Responda si es cierto o falso (y corrija, si procede):

- a) Tanto las empresas con ganancias extremadamente altas como las empresas con ganancias extremadamente bajas en 1999, tendieron a tener ganancias menos extremas en el 2000.
- b) Lo anterior indica, aunque no demuestra, que algún factor (quizás una política de fijación de impuestos progresista o una revisión conservadora de parte de los ejecutivos de administración,...) provocó que las ganancias fueran menos extremas en el años 2000 que en el año 1999.
- c) Lo anterior muestra, entre otras cuestiones, cuán difícil es permanecer cerca de la cima.
- d) La dispersión es simétrica con respecto a la recta de 45° donde $G_{2000} = G_{1999}$, por consiguiente, para una empresa cuya ganancia en 1999 fue del 11%, la mejor predicción para su ganancia en el 2000 es también el 11%.

7) SELECCIONANDO CASOS. Es frecuente seleccionar a los pacientes incluidos en un EC siguiendo diferentes criterios. Por ejemplo, si el fenómeno principal en estudio es la presión arterial sistólica (PAS), se le requiere al paciente que supere, en dos desviaciones típicas, el promedio de PAS en los pacientes sanos ($\mu + 2\sigma$).

Busque por internet información sobre el fenómeno de regresión a la media (p.e., http://en.wikipedia.org/wiki/Regression_toward_the_mean) ¿Cuánto cabe esperar que mejore espontáneamente la PAS? Asuma que todos los pacientes tienen exactamente el valor $\mu + 2\sigma$; y que las dispersiones de ambas determinaciones son iguales. Calcule el valor esperado en una segunda determinación. Asuma que la correlación entre determinaciones repetidas de PAS vale 0.75.

8) Proponga algún ejemplo próximo de regresión a la media. Explique, con la ayuda de gráficos y números qué pasa.

9) Invente unos datos con 3 variables (Y, respuesta; X, intervención; Z, condición) dicotómicas en los que haya sesgo de selección.

10) Invente una situación hipotética experimental (un ensayo clínico) con 3 variables (Y, respuesta; X, intervención; Z, condición) en la que pueda existir sesgo de selección. Dibuje el DAG.

11) Invente una situación hipotética observacional con 3 variables (Y, respuesta; X, intervención; Z, condición) en la que pueda existir sesgo de selección. Dibuje el DAG.

12) LO OBSERVADO Y LO QUE HABRÍAMOS OBSERVADO SI... El protocolo de intervención de una neoplasia aconseja que los casos de nivel I (50%) deben seguir una pauta de cirugía más radioterapia (C), mientras que los de nivel II, quimioterapia más radioterapia (Q). En cierto centro, una tercera parte de los 90 casos observados no siguen las recomendaciones y son finalmente tratados con la otra opción. Se observa que se curan ($Y=0$) 27 de los tratados con C y 45 del total de casos.

- Calcule el OR observado entre X (C/Q) e Y (0/1)
- Sabiendo que ambos tratamientos tienen la misma eficacia, reconstruya el diagrama de árbol con las probabilidades condicionadas que faltan.
- Utilice el IPW para estimar el efecto de cambiar C por Q

13) Tras sufrir un Ictus o Accidente Vascular Cerebral (AVC), los pacientes tardan un tiempo muy variable en llegar al hospital, que condiciona tanto la evolución como las intervenciones disponibles y que depende del Entorno, sea Rural, $P(E=R)=1/3$, o Urbano, $P(E=U)=2/3$. Por ejemplo, en el AVC isquémico (no hemorrágico) se aconseja intentar disolver el coágulo mediante la administración de agentes tromBolíticos ($X=B$ frente a $X=No$) sólo a los pacientes que llegan hasta 3 horas desde el inicio del AVC. En R, la proporción de casos tratados con B es $1/5$, mientras que en U es de $4/5$. En ambos entornos, el efecto de B frente a NO en la evolución (positiva: $Y=1$, frente $Y=0$) tiene un $OR=4$.

- Reconstruya el árbol sabiendo que el total de casos es 45.

E_R	Y_0	Y_1
X_B		
X_{NO}		

E_U	Y_0	Y_1
X_B		
X_{NO}		

- Estime el efecto de cambiar B por NO sin ajustar.
- Ahora deseamos saber cuál sería el efecto observado si tratáramos a todos los pacientes con la misma opción, sea B, sea N: ajuste por IPW.
- Interprete conjuntamente las 2 preguntas anteriores.
- ¿Tienen la misma evolución los que viven en E_R que en E_U ? Hágalo sin ajustar.
- Ahora ajuste por IPW, asumiendo que deseamos conocer qué pasaría si situamos a todos los pacientes en el entorno R respecto a si los situamos a todos en U (o, si conseguimos que los tiempos de acceso al hospital sean los mismos en R que en U).
- Interprete conjuntamente las 2 preguntas anteriores.

14) Reciben el nombre de micro-organismos 'oportunistas' aquellos que suelen ser inocuos en personas sanas pero son capaces de provocar enfermedades en pacientes debilitados. En Hospitales, suelen ser los responsables de las llamadas infecciones

nosocomiales (IN) que aparecen durante la estancia (E=Larga frente a Corta) en el hospital. De hecho, un indicador de calidad muy relevante es precisamente la proporción de pacientes ingresados que desarrollan IN. Los responsables de calidad del centro comparan entre servicios (A/B) la proporción IN para todos los pacientes que están ingresados más de 1 semana (Larga). Vd. sabe que la estancia varía entre servicios. Y sabe que seleccionar a los que más están podría provocar un sesgo de selección.

- Dibuje el DAG de las variables servicio (S=A,B), longitud estancia (E=C, L) e IN (S,N) y explique porqué seleccionar sólo a los que tienen estancia larga puede provocar sesgo de selección.
- De hecho, ambos servicios atienden sólo 2 enfermedades (Códigos C=23, 24), en distinta proporción, que deberían anular la relación entre S Y E. Dibuje un nuevo DAG y proponga un (conjunto de) análisis que permita comparar la calidad de los servicios en cuanto a IN.

15) UNA PREVENCIÓN PELIGROSA. Para prevenir que los pacientes ingresados en un hospital desarrollen complicaciones (Y=S vs Y=N), suele hacerse una intervención de enfermería (E=S vs E=N) consistente en movilización, masaje y otros componentes. Una escala pronóstica permite clasificar a los pacientes en 2 grupos según el mayor o menor riesgo (pronóstico malo, P=M vs pronóstico bueno, P=B) a presentar complicaciones. Para valorar el efecto de la intervención de enfermería (S:N) decidimos hacer un EC aleatorizando con una razón de aleatorización (S:N) que depende del bloque pronóstico: 9:1 en P=M, pero 1:9 en P=B. La siguiente tabla muestra los resultados para los pacientes que han permanecido ingresados más de 4 días.

P=M	Y=S	Y=N	P=B	Y=S	Y=N
E=S	480	240	E=S	16	38
E=N	80	80	E=N	248	1184

- Calcule e interprete el efecto de la intervención en ambos grupos de P.

Alertados por un posible sesgo de selección, los investigadores deducen de otra información externa que la proporción de pacientes que permanecen ingresados más de 4 días es un 80%, excepto para aquellos casos que sí han recibido la intervención (E=S) y no han presentado complicaciones (Y=N), que es de un 20%.

- Calcule otra vez el efecto de la intervención, en ambos grupos de pronóstico, teniendo en cuenta mediante IPW esta selección de casos.
- Repita (1) y (2), pero sin condicionar por pronóstico. Discuta la conveniencia o no de realizar este ajuste. ¿Qué nombre recibiría esta situación?

16) Cierta empresa procesa un producto tóxico (T) del que se sospecha que provoca varias enfermedades (E) que son, por otro lado, frecuentes en el entorno de la empresa. Instado por el empresario, el médico recién llegado estudia a los 26 trabajadores para ver si tienen o no la enfermedad y averigua si estuvieron o no expuestos. Observa un total de 23 casos $\neg E$, 10 casos $\neg T$ y 2 casos que son, a la vez, E y T.

a. Reconstruya la tabla 2x2, calcule el OR (cuyo IC_{95%} es 0.10 a 16.34) e interprete. (3p)

	E	No E	Total
T			
No T			
Total			

b. Al empresario no le gustan los resultados y busca en los archivos de la empresa los trabajadores que había tenido. Tras consultar con ellos, recuerdan que los expuestos habían sido muy vigilados y por ello, aquellos que resultaban afectados por la enfermedad (ET) tenían gran tendencia a cambiar de entorno, por lo que, desde sus inicios, aproximadamente 5 de cada 6 del grupo ET dejaron la empresa, mientras que en el resto de grupos, (E $\neg T$, $\neg E$ T, $\neg E$ $\neg T$) lo han hecho 1 de cada 2. Usando el IPW calcule de nuevo el OR resultante (cuyo IC_{95%} aproximado va de 1.24 a 12.04). (4p)

c. ¿Qué nombre recibe esta situación? ¿Por qué? (2p)

d. Interprete globalmente los resultados. (4p)

e. ¿Por qué están más protegidos ante esta situación aquellos estudios que seleccionan a los pacientes por criterios iniciales (cohortes, ensayos clínicos) que los que los seleccionan por criterios finales (casos y controles)? (2p)

17) HOMOGENEIDAD DEL EFECTO. Se desea estudiar si cierto analgésico T tiene el mismo efecto en ayunas (A) o no (N). Un EC selecciona 20 casos en cada condición de ayuno, asigna la mitad al analgésico o al placebo P y mide el dolor al cabo de 30' en una escala Y que va de 0 a 100 y que tiene una varianza en los casos reclutados de 32 puntos². Se decide contrastar el efecto de T frente a P estimando $E(Y_T) - E(Y_P)$.

a. Calcule la varianza del estimador del efecto.

Se decide contrastar la heterogeneidad del efecto de T frente a P en los subgrupos de A y N estimando $[E(Y_{TA}) - E(Y_{PA})] - [E(Y_{TN}) - E(Y_{PN})]$.

b. Calcule la varianza del estimador de la heterogeneidad del efecto.

c. Si deseamos poder demostrar un efecto de T frente a P de $\Delta=5$ puntos al nivel alfa 5% bilateral, ¿de qué potencia disponemos?

- d. Si deseamos poder detectar una heterogeneidad del efecto de T frente a P de $\Delta^H=5$ puntos al nivel alfa 5% bilateral, ¿de qué potencia disponemos?
- e. Intente relacionar Δ^H y Δ . interprete globalmente los resultados: ¿Qué opina sobre contrastar la heterogeneidad del efecto?
- f. Teniendo en cuenta que, para tener utilidad práctica, Δ^H debería ser menor que Δ , interprete globalmente los resultados: ¿Qué opina sobre contrastar la heterogeneidad del efecto?

18) PREMISAS SOBRE VARIABLES NO OBSERVABLES. Entre las premisas de suficiencia del modelo, ausencia de variables confusoras y aleatorización, existen 2 implicaciones. ¿Cuáles son? Pista: clasifique la variables V no observables en aquellas que influyen en la respuesta (V_c) y aquellas que no (V_o).

19) Vd. sabe que Vd. cumplirá con las recomendaciones del médico. Estime el efecto causal promedio en los cumplidores con los datos de la siguiente tabla.

Allocated treatment	Treatment received	Complier	Outcome (stroke)	Cases
1	1	C	0	1381
1	1	C	1	37
1	0	N	0	341
1	0	N	1	103
0	0		0	165
0	0		1	222

20) Estime efecto ITT y efecto casual promedio en los cumplidores para una respuesta numérica Y (escala de dolor entre 0, no dolor y 20, dolor máximo)

Allocated treatment	Treatment received	Complier	Outcome mean	Cases
1	1	C	6	80
1	0	N	12	20
0	0		13	100