

II Jornades de
consultoria
estadística i
software

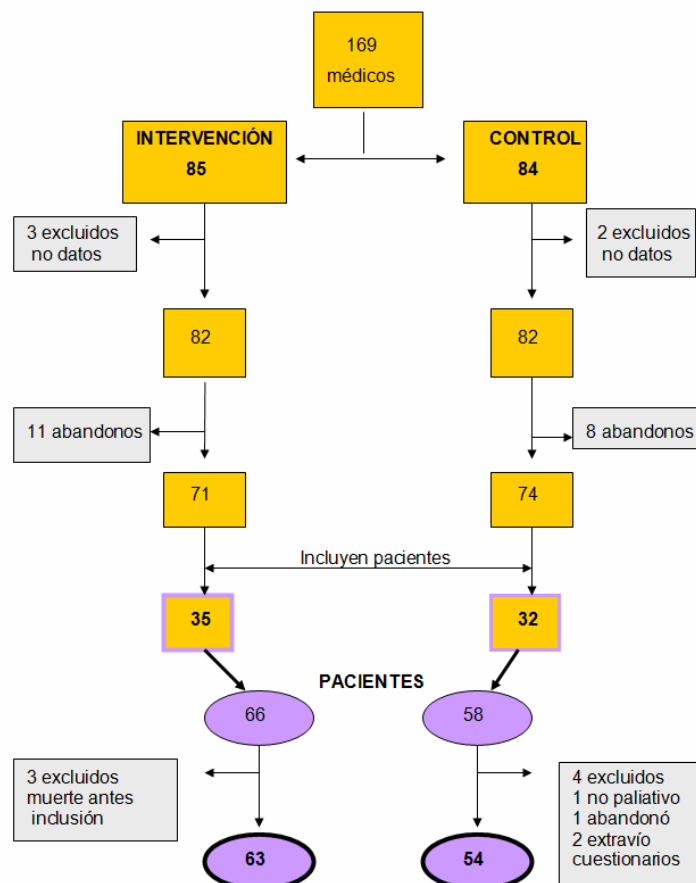
Imputació de dades en l'avaluació de formació en línia en cures pal·liatives en Metges d'Atenció Primaria

Santiago Pérez-Hoyos
Unitat Suport Metodològic a l'Investigació Biomèdica (USMIB)
Vall d'Hebron Institut de Recerca (VHIR)
Planta 13a. Àrea MaternoInfantil
santi.perezhoyos@vhir.org

Antecedentes

- Grupo de Médicos de AP con habilidades en análisis
 - Estudio para evaluar los efectos de la formación “on line” en los pacientes de cuidados paliativos
 - Uso de Escalas de dolor, cuidados paliativos, calidad de vida
 - Cada escala es combinación de ítems y tiene diferentes dimensiones
 - Información faltante en un ítem invalida el cálculo de la dimensión de la escala y la combinación podía invalidar numerosas observaciones
1. Si merece la pena imputar los datos ausentes de los ítems de los cuestionarios
 2. Qué tipo de análisis multivariante es el más adecuado a la respuesta principal y secundaria: análisis de la covarianza, RLM? Bonferroni?

- Experiencia en imputación de fechas de infección en estudios de VIH
- No experiencia en variables cualitativas
- Equipo de investigación con buena interlocución
- Búsqueda de solución que se presenta a continuación



Cuestionario Breve del dolor, (1 y 2)

11 ítems , dos dimensiones (impacto e intensidad) y 81% casos completos

**Escala de cuidados paliativos, versión médico (1 y 2)
 Escala de cuidados paliativos, versión paciente (1 y 2)**

11 ítems y 74% casos completos

Escala calidad vida Rotterdam (1 y 2)

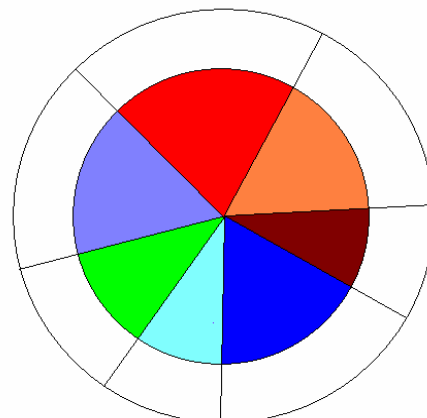
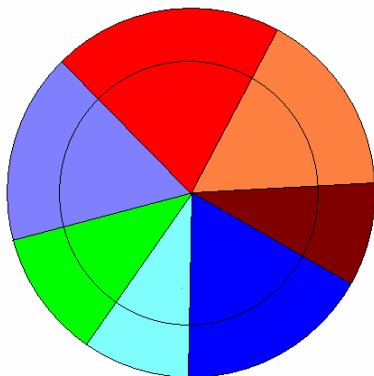
10 ítems y 81 % casos completos

Cuestionario Satisfacción cuidador SERVQUAD

5 dimensiones apenas missings

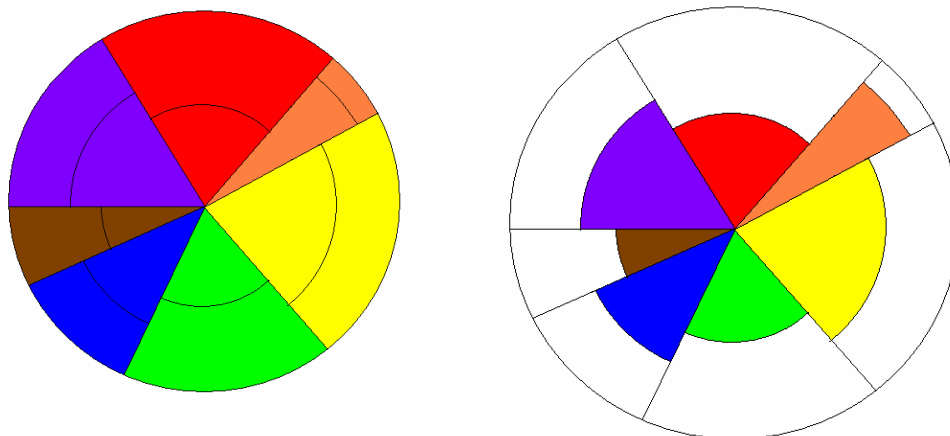
Test de conocimiento (0, 4 y 18 meses) Test de conocimiento 53% 3 medidas

**MCAR (missing completely at random)
 Los datos perdidos son totalmente al azar**



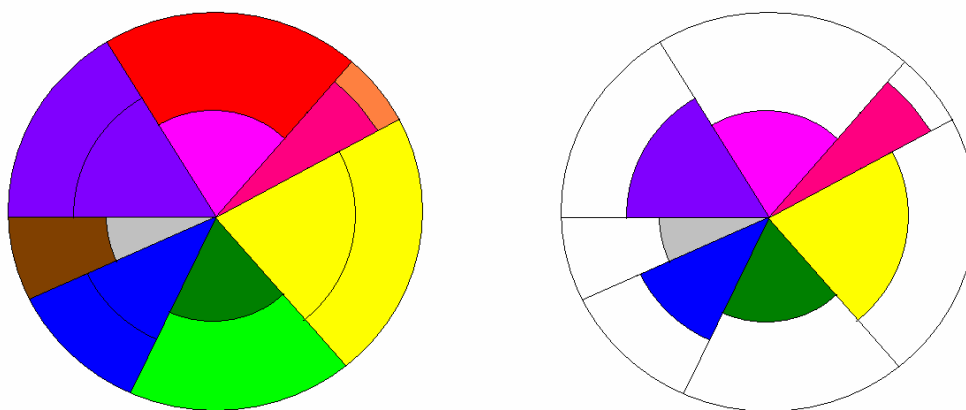
MAR (missing at random)

Los valores de los datos perdidos no dependen de la información faltante, aunque puede hacerlo de la observada.

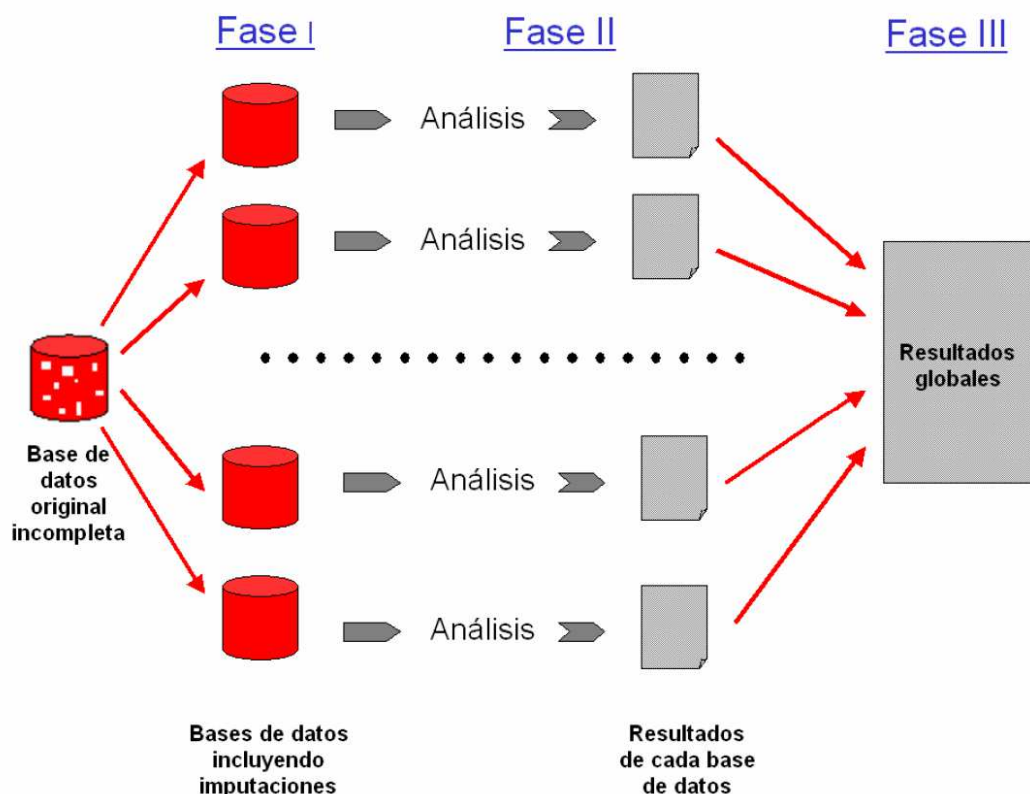


MNAR (missing not at random)

La probabilidad del dato faltante depende de la información no observada



- Imputación por media
- Imputación Cold-Deck (predicción modelo externo)
- Imputación Hot-Deck (asignación individuo semejante)
- Imputación por regresión
- Imputación por Máxima-verosimilitud (Algoritmo EM)
- Imputación múltiple



Generar múltiples conjuntos de datos imputados

Las variables x_1, \dots, x_k tienen datos faltantes

Eliminar casos sin información

Inicializar los valores perdidos al azar

Regresar x_1 en x_2, \dots, x_k según tipo x_1 (Gaussiana, logística, ordinal)

Reemplazar valores perdidos x_1 con predicción en x_1

Repetir para x_2, \dots, x_k sobre otras x 's (ciclo 1)

Repetir sobre 10 ciclos

Repetir el proceso completo m veces (5)

Se generan m conjuntos de datos imputados con todos los datos

Combinación de las estimaciones

Se combinan las m estimaciones y la matriz varianza-covarianza siguiendo las reglas de Rubin.

Se incorpora la variabilidad de cada imputación con la variabilidad entre imputaciones

$$\hat{\theta} = \frac{1}{m} \sum_{j=1}^m \hat{\theta}_j$$

$$\text{var}(\hat{\theta}) = \frac{1}{m} \sum_{j=1}^m W_j + \left(1 + \frac{1}{m}\right) \left(\frac{1}{m-1}\right) \sum_{j=1}^m (\hat{\theta} - \hat{\theta}_j)^2$$


```
ice ecyme1 ecyme2 ecyme3 ecyme4 ecyme5 ecyme6 ecyme7 ecyme8 ecyme9 ecyme11 ///
f_ecyme1 f_ecyme2 f_ecyme3 f_ecyme4 f_ecyme5 f_ecyme6 f_ecyme7 f_ecyme8 f_ecyme9 f_ecyme11 ///
ecppa1 ecppa2 ecppa3 ecppa4 ecppa5 ecppa6 ecppa7 ecppa8 ecppa9 ecppa11 ///
f_ecppa1 f_ecppa2 f_ecppa3 f_ecppa4 f_ecppa5 f_ecppa6 f_ecppa7 f_ecppa8 f_ecppa9 f_ecppa11 ///
pedad psexo pdolor ptiempo pkps pconoce ///
```

, saving("\$dta/imputa_ecp.dta",replace) m(5) clear cmd(ologit) seed (199239123)

```
ice psexo pedad pdolor ptiempod pkps pconoce fisica fisica1 psicolog activida global fatiga dolor gastroin ///
quimiote diasentr f_fisica f_psicolog f_activida f_global f_fatiga f_dolor f_gastroin f_quimiote ///
```

, saving("\$dta/imputa_rscl",replace) m(5) clear seed (199239123) ///

cmd(fisica psicolog activida global fatiga dolor gastroin quimiote diasentr f_fisica f_psicolog f_activida f_global f_fatiga f_dolor f_gastroin f_quimiote:regress, psexo pedad pdolor ptiempod pkps pconoce:ologit)

Cuestionario Breve del dolor, (1 y 2)

11 ítems , dos dimensiones (impacto e intensidad) y 81% casos completos

Signo-rango Wilcoxon, T-test compara grupos, Modelo lineal Mixto

**Escala de cuidados paliativos,
versión médico (1 y 2)
Escala de cuidados paliativos,
versión paciente (1 y 2)**

11 ítems y 74% casos completos

Concordancia, Kappa, Spearman, Test Homogeneidad

**Escala calidad vida Rotterdam
(1 y 2)**

10 ítems y 81 % casos completos

Signo-rango Wilcoxon, T-test compara grupos, Modelo lineal Mixto

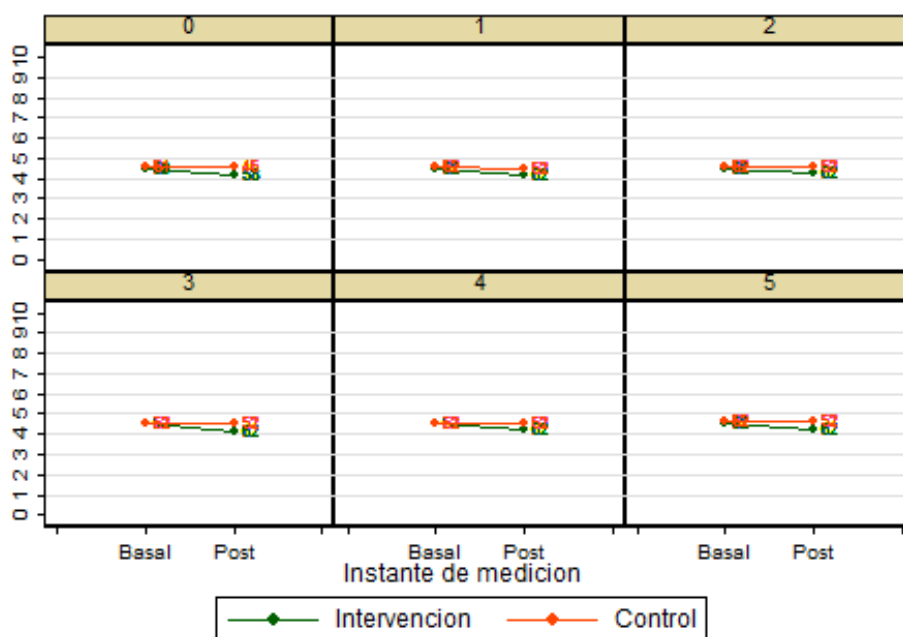
**Cuestionario Satisfacción
cuidador SERVQUAD**

5 dimensiones apenas missings

T-test compara grupos,

Test de conocimiento (0, 4 y 18 meses) Test de conocimiento 53% 3 medidas

Modelo lineal Mixto



Graphs by imputation number

xi

<u>Intensidad</u>	<u>Primera medicación</u>		<u>Segunda medicación</u>	
	<u>Original Media</u> [IC 95%]	<u>Imputación Media</u> [IC 95%]	<u>Original Media</u> [IC 95%]	<u>Imputación Media</u> [IC 95%]
<u>Intervención</u>	4.51 [3.58; 4.67]	4.49 [3.94; 5.05]	4.12 [3.58; 4.67]	4.25 [3.70; 4.80]
<u>Control</u>	4.58 [4.00; 5.16]	4.59 [4.01 ; 5.18]	4.57 [3.91; 5.22]	4.64 [3.70; 4.80]
<u>Diferencia</u>	-0.070 [-0.870; .730]	-0.099 [-0.900; 0.703]	-0.446 [-1.28; .392]	-0.383 [-1.220 ; 0.453]
<u>Valor p</u>	0.863	0.806	0.294	0.346

xi: mim: xtmixed intensidad i.grupo*i.momento || identifica:,mle

Datos originales	Estimación	IC.95%	Valor p	Diferencia con el grupo de referencia
Primera medición grupo intervención (Referencia)	4.40	[3.87; 4.93]		0
Efecto de la segunda medición	-0.22	[-0.59; 0.15]	0.246	-0.22
Efecto del grupo control	0.13	[-0.65; 0.91]	0.740	+0.13
Interacción entre la segunda medición y el grupo control	0.28	[-0.27; 0.83]	0.319	+0.19= 0.28-0.22+0.13
Diferencia segunda medición en el grupo control	0.06	[-0.35; 0.47]	0.768	+0.06= -0.22+0.28
Diferencia grupo intervención y grupo control en segunda medición	0.41	[-0.38; 1.20]	0.3068	+0.41= 0.28+0.13

Datos imputados				Diferencia con el grupo de referencia
Primera medición grupo intervención (Referencia)	4.49	[3.96;5.02]		0
Efecto de la segunda medición	-0.24	[-0.65 ; 0.17]	0.247	-0.24
Efecto del grupo control	0.10	[-0.69 ; 0.89]	0.808	+0.14
Interacción entre la segunda medición y el grupo control	0.28	[-0.39 ; 0.96]	0.403	+0.16= 0.28-0.24+0.10
Diferencia entre segunda y primera medición en el grupo control	0.04	[-0.50 ; 0.59]	0.352	+0.04= -0.24+0.28
Diferencia entre grupo intervención y grupo control en segunda medición	0.38	[-0.43 ; 1.20]	0.356	+0.38= 0.28+0.10

Clinical Effectiveness of Online Training in Palliative Care of Primary Care Physicians

Marta Pelayo-Alvarez, MD,¹ Santiago Perez-Hoyos, PhD,² and Yolanda Agra-Varela, PhD³

TABLE 2. GROUP COMPARISON

Original data Imputed data	Value	95% CI	P	Difference with reference group
Pain intensity				
Time 1 in intervention (reference)	4.40	[3.87-4.93]	—	0
	4.49	[3.96-5.02]		0
Effect in time 2	-0.22	[-0.59-0.15]	0.246	-0.22
	-0.24	[-0.65-0.17]	0.247	-0.24
Effect in control	0.13	[-0.65-0.91]	0.740	0.13
	0.10	[-0.69-0.89]	0.808	0.14
Interaction between time 2 and control	0.28	[-0.27-0.83]	0.319	0.19
	0.28	[-0.39-0.96]	0.403	0.16
Difference between time 2 and time 1 in control	0.06	[-0.35; 0.47]	0.768	0.06
	0.04	[-0.50-0.59]	0.352	0.04
Difference of intervention and control in time 2	0.41	[-0.38-1.20]	0.307	0.41
	0.38	[-0.43-1.20]	0.356	0.38
Pain impact				
Time 1 in intervention (reference)	6.32	[5.56-7.07]	—	0
	6.41	[5.68-7.15]		0
Effect in time 2	-0.18	[-0.67-0.30]	0.572	-0.18
	-0.49	[-1.04-0.07]	0.084	-0.49
Effect in control	0.32	[-0.79-1.42]	0.460	0.32
	0.28	[-0.82-1.37]	0.617	0.28
Interaction between time 2 and control	0.23	[-0.49-0.94]	0.532	0.37
	0.46	[-0.34-1.26]	0.259	0.25
Difference between time 2 and time 1 in control	0.05	[-0.48-0.56]	0.871	0.05
	-0.03	[-0.71-0.66]	0.941	-0.03
Difference of intervention and control in time 2	0.55	[-0.58-1.67]	0.342	0.55
	0.38	[-0.36-1.84]	0.039	0.74

Brief Pain Inventory scale is 1 to 11.
CI, confidence interval.

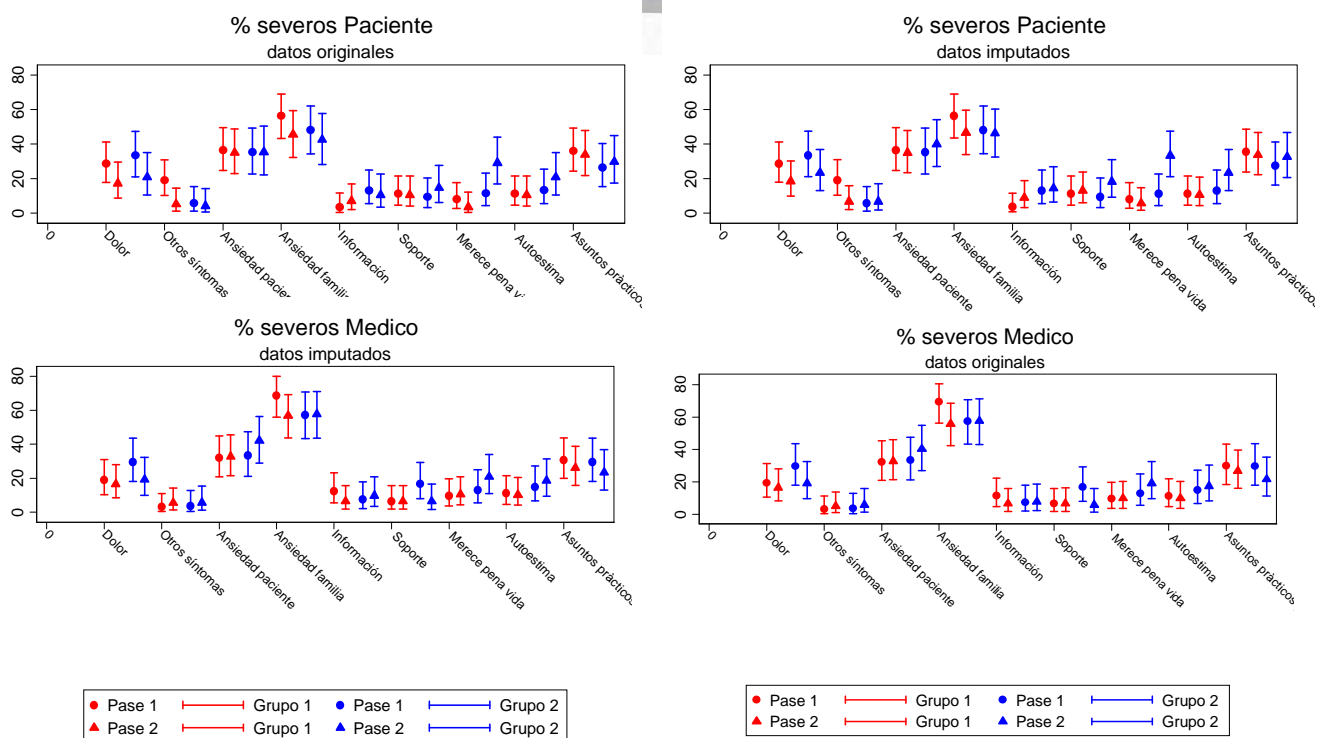
Datos originales

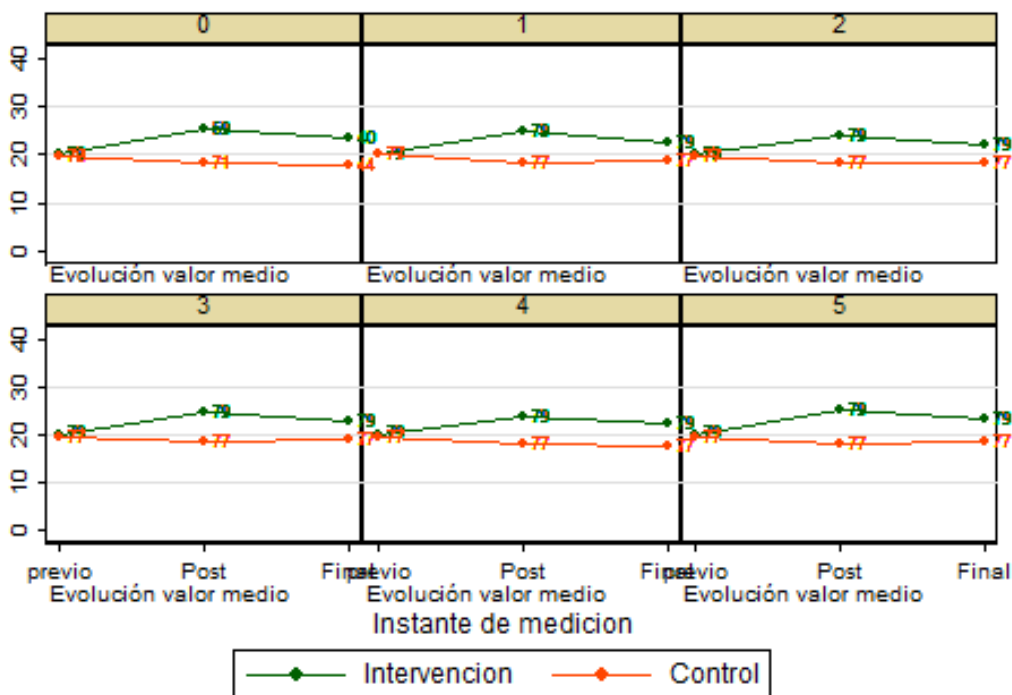
	Item ECP	Grupo	Pase de la medición	% severos Paciente	Lim. Inf.	Lim.Sup	% severos Medico	Lim. Inf.	Lim.Sup	Kappa	Lim. Inf.	Lim.Sup	Cor. Spearman	% acuerdo
1	Dolor	1	1	28.57	17.89	41.35	19.35	10.42	31.37	.51	.34	.68	.655	85.89
2	Dolor	2	1	33.33	21.09	47.47	29.63	17.98	43.61	.54	.36	.71	.699	85.65
3	Dolor	1	2	17.24	8.59	29.43	16.39	8.15	28.09	.57	.39	.74	.655	88.79
4	Dolor	2	2	20.83	10.47	34.99	19.23	9.63	32.53	.59	.41	.77	.724	89.06

mim ,cat(combine) byvar est(r(kappa)) se(r(se)) :kap ecppa ecpme', absolute wgt(w)

Datos Imputados

	Item ECP	Grupo	Pase de la medición	% severos Paciente	Lim. Inf.	Lim.Sup	% severos Medico	Lim. Inf.	Lim.Sup	Kappa	Lim. Inf.	Lim.Sup	Cor. Spearman	% acuerdo
1	Dolor	1	1	28.57	17.89	41.35	19.05	10.25	30.91	.5	.34	.67	.65	85.63
2	Dolor	2	1	33.33	21.09	47.47	29.63	17.98	43.61	.54	.36	.71	.699	85.65
3	Dolor	1	2	18.41	9.77	30.18	16.51	8.35	27.99	.53	.37	.7	.632	87.94
4	Dolor	2	2	23.33	12.93	36.8	19.26	9.81	32.27	.55	.38	.72	.693	87.87





Graphs by imputation number

Tabla RESUMEN TEST Datos originales

Variable	Diferencia aritmética	(95% Intervalo Conf.)	P-valor	P-valor Global
<i>Parte fija</i>				
Test inicial grupo intervención	20.24	(19.50; 20.98)	.	-
Diferencia inicial grupo control	-0.44	(-1.50; 0.62)	0.4153	-
Diferencia a los 4 meses con test inicial en grupo intervención	5.08	(4.13; 6.03)	0.0000	-
Diferencia a los 18 meses con test inicial en grupo intervención	2.98	(1.88; 4.07)	0.0000	-
Test inicial grupo control	19.80	(19.03; 20.56)	.	-
Diferencia a los 4 meses con test inicial en grupo control	-1.76	(-2.67; -0.86)	0.0001	-
Diferencia a los 18 meses con test inicial en grupo control	-1.76	(-2.83; -0.70)	0.0012	-

Tabla RESUMEN TEST Datos Imputados

Variable	Diferencia aritmética	(95% Intervalo Conf.)	P-valor	P-valor Global
<i>Parte fija</i>				
Test inicial grupo intervención	20.24	(19.50; 20.98)	.	-
Diferencia inicial grupo control	-0.492	(-1.80; 0.82)	0.462	-
Diferencia a los 4 meses con test inicial en grupo intervención	4.23	(2.25; 6.21)	0.001	-
Diferencia a los 18 meses con test inicial en grupo intervención	2.32	(.88; 3.76)	0.003	-
Test inicial grupo control	19.75	(18.79; 20.70)	.	-
Diferencia a los 4 meses con test inicial en grupo control	-1.46	(-2.54; -0.38)	0.008	-
Diferencia a los 18 meses con test inicial en grupo control	-1.28	(-2.79; 0.23)	0.092	-

- En ocasiones las consultorías nos plantean problemas nuevos
- Para los usuarios los datos son valiosos, ¿Porque perderlos?
- Algunas veces no se pueden ajustar modelos sensatos sin imputación, pero hay que hacerlo con sensatez
- Se puede efectuar bajo la hipòtesis MAR pero a menudo no se puede testar
- Puede ser costoso el cálculo de las imputaciones
- Permite mejorar la precisión y la estimación cuando hay datos faltantes.
- Stata, R y SAS permiten efectuar fácilmente imputaciones

La formación online en CP dirigida a MAP disminuyó las puntuaciones severas de síntomas y ansiedad familiar. incrementó la escala global de calidad de vida del paciente, sin evidenciar diferencias significativas en el resto de síntomas y calidad de vida del paciente, y en la satisfacción del cuidador principal al ser comparada con la formación tradicional o no formación específica. Los MAP incrementaron significativamente los conocimientos a corto y largo plazo.

Gràcies

Santiago Pérez-Hoyos
Unitat Suport Metodològic a l'Investigació Biomedica
(USMIB) Vall d'Hebron Institut de Recerca (VHIR)
Planta 13a. Area MaternoInfantil
santi.perezhoyos@vhir.org