

Màster Interuniversitari en Estadística i Investigació Operativa

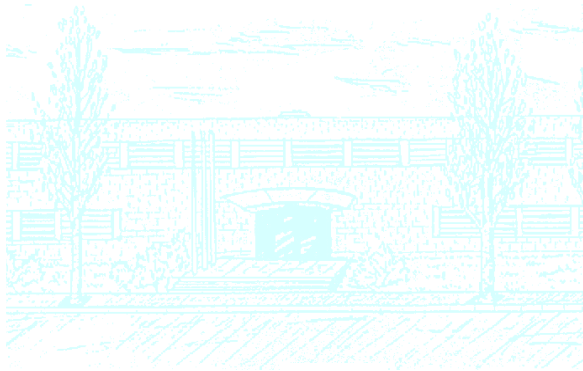
Títol: Anàlisi de mesures repetides en assaigs clínics:
Estudi de l'efecte hipotèrmic de la Citicolina.
Autor: Jordi Blanch i Font.

Director: Julio Secades, Erik Cobo Valeri i Josep Anton
Sánchez Espigares.

Departament: Departament d'Estadística i Investigació
Operativa

Universitat: Universitat Politècnica de Catalunya.

Convocatòria:



Facultat de Matemàtiques
i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA



UNIVERSITAT DE BARCELONA



Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Tesi de màster

**Anàlisi de mesures repetides en assaigs
clínic: Estudi de l'efecte hipotèrmic de la
Citicolina.**

Jordi Blanch i Font

Director: Julio Secades
Erik Cobo Valeri
Josep Anton Sánchez Espigares

Departament d'Estadística i Investigació Operativa

Resum

Paraules clau: Accidents vasculars cerebrals, hipotèrmia, anàlisi de mesures repetides, valor pronòstic, Citolina, efectes en la temperatura.

MSC2000: 200062-07 Data analysis
62J05 Linear regression
62J10 Analysis of variance and covariance
62J12 Generalized linear models
62P10 Applications to biology and medical sciences

Les malalties vasculars cerebrals són la tercera causa de mort en els països desenvolupats i una de les principals causes de discapacitat i de seqüeles neurològiques. Hi han molts assaigs clínics que han tractat de reduir la mortalitat i/o la gravetat de les seqüeles.

Analitzarem l'evolució de la temperatura durant l'assaig per a estimar si el tractament amb Citolina modifica l'evolució. També estudiarem el valor pronòstic de la temperatura basal sobre la mortalitat i la gravetat de les seqüeles. L'assaig estudiat té 899 pacients dividits en: controls i tractats. Els pacients se'ls controla durant sis setmanes de tractament i sis més de seguiment. Es va concloure que no hi havia suficient evidència per a establir un efecte sobre les seqüeles.

Utilitzarem un mètode diferent per a respondre cada problema. Per a l'efecte de la Citolina hem construït el millor model per a l'evolució sense el tractament i, posteriorment, s'ha inclòs el tractament per a estimar el seu efecte. Per a la capacitat predictiva de la temperatura basal, hem utilitzat un model logístic per la mortalitat i un logístic proporcional ordinal per a les seqüeles.

El model per a l'evolució de la temperatura és una aproximació quadràtica del temps. Per ajustar millor, s'han inclòs uns efectes aleatoris sobre cada terme i una funció de variància per a l'heterocedasticitat de les dades. L'estimació de l'efecte del tractament en el terme lineal és $LT = 0,002181$ ($\sigma_{LT} = 0,002122$) i el quadràtic és $QT = -0,000022$ ($\sigma_{QT} = 0,000022$).

En el model per a la mortalitat, s'estima que l'augment d'un grau de la temperatura basal suposa un *OR* ajusta de 0,9149 ($IC_{95\%} = 0,6749 - 1,2401$) sobre la mort del pacient. Per la gravetat de les seqüeles, l'augment d'un grau de la temperatura basal té associat un *OR* ajustat de 0,9134 ($IC_{95\%} = 0,7339 - 1,1358$).

El model per a l'evolució de les temperatures és complet amb les variables disponibles. Si el model és correcte, no s'ha demostrat que hi hagi suficient evidència per dir que l'efecte del tractament és significatiu. Tampoc s'ha trobat que existeixi suficient evidència per a poder afirmar que la temperatura basal ajuda a predir la mortalitat o la gravetat de les seqüeles.

Abstract

Keywords: Cerebral vascular diseases, hypothermia, Repeated measures analysis, prognostic value.

MSC2000: 200062-07 Data analysis
62J05 Linear regression
62J10 Analysis of variance and covariance
62J12 Generalized linear models
62P10 Applications to biology and medical sciences

Cerebral vascular diseases are the third leading cause of death in developed countries, one of the leading causes of disability and neurological sequelae. There are many clinical trials have sought to reduce mortality and/or severity of sequelae.

We will focus on the body temperature behaviour of the [3] clinical trial. We analyze the temperature evolution throughout the trial to estimate whether the treatment altered the treatates development. We also quantify the baseline temperature prognostic value on mortality and the sequelae severity. The trial had 899 patients divided into: controls and treated. Patients were followed-up six weeks with treatment and six more without. The trial concluded that there was not sufficient evidence to accept that treatment was effective.

We apply a different method to answer each problem. For the temperature evolution, we build the best model for the evolution without treatment and then treatment was included to estimate its effect. For the relationship between baseline temperature and mortality, we used a logistic model. Finally, for the sequelae severity, we used ordinal logistic regression model with the odds ratio (*OR*) proportionality.

The model for the temperature evolution is a quadratic approximation of time. We included random effects to adjust better the evolution of each patient. Moreover, the variance decreased through the study is modelled with a variance function. The interaction between the treatment and the time is $LT = 0.002181$ ($\sigma_{LT} = 0.002122$) and the interaction between the treatment and the square of time is $QT = -0.000022$ ($\sigma_{QT} = 0.000022$).

In the model for mortality, we estimated that a one degree increase of baseline temperature have a adjusts *OR* of 0.9149 ($CI(OR, 95\%) = 0.6749 - 1, 2401$) on the death of the patient. For the sequelae gravity, increasing a one degree of baseline temperature has an associated adjusted *OR* of 0.9134 ($CI(OR, 95\%) = 0.7339 - 1.1358$).

The model for the temperature evolution is complete with study variables. Assuming that the model is correct, there is note enough evidence to say that the treatment effect is significant. We found that there is not enough evidence to say that for the basal temperature is a good prognostic variable of mortality or sequelae severity.

Índex general

Capítol 1. Introducció	1
1. Malalties vasculars cerebrals	1
2. Objectius i Hipòtesis	3
Capítol 2. Materials i Mètodes	5
1. Població d'estudi	5
2. Mètodes estadístics	8
Capítol 3. Resultats	19
1. Descriptiva general	19
2. Hipotèrmia induïda pel tractament	21
3. La mort en funció de la temperatura basal	33
4. La temperatura basal com a factor pronòstic	38
Capítol 4. Discussió	43
1. Resum dels resultats	43
2. Models estadístics	43
3. Explicació de l'efecte observat	44
4. Comparació amb altres articles	44
5. Limitacions	46
6. Recerca futura	46
Apèndix A. Construcció del model sense tractament	49
1. Model amb pacients amb totes les mesures	57
Apèndix B. Taules de la linealitat dels logits de morir	59
1. Taules dels logits	59
2. Assumpció de linealitat dels logits de l'edat	61
Apèndix C. Taules de la linealitat dels cumulative logits de tenir seqüeles	63
Apèndix D. Codi R	65
1. Descriptiva General	65
2. Hipotèrmia induïda pel tractament	66
3. La mort en funció de la temperatura basal	68
4. La temperatura basal com a factor pronòstic	72
5. Annex: Construcció del model sense tractament	79
6. Annex: Model amb pacients amb el tractament complet	82

Apèndix. Bibliografia

85

Capítol 1

Introducció

S'han realitzat diversos assaigs clínics per verificar la eficàcia de la citicolina davant del placebo per a tractar pacients que han patit una malaltia vascular cerebral isquèmica ([2], [4], [3] i [5]). Aquests assaigs només van demostrar el benefici clínic per a dosis altes de citicolina. Però les dosis altes augmenten el risc de vertigen i de danys cerebrals. Es va desestimar com a medicació per a les malalties vasculars cerebrals.

Aquests estudis es van mostrar comportaments diferents entre els tractats i els controls. Respecte la temperatura corporal, el grup tractat va tenir una disminució de la temperatura més accentuat que els pacients controls. Aquest treball pretén demostrar aquesta hipòtesis prèvia.

1. Malalties vasculars cerebrals

Les malalties o accidents vasculars cerebrals (AVC) són un dels motius més freqüents d'assistència mèdica urgent al nostre medi. Són trastorns clínics derivats d'una aportació insuficient de sang al sistema nerviós central (SNC).

Les malalties vasculars cerebrals són la tercera causa de mortalitat als països desenvolupats i la primera causa de discapacitat i de seqüeles neurològiques a la població adulta. Aquestes seqüeles produeixen un gran impacte social i econòmic.

L'incidència de les malalties vasculars cerebrals augmenta amb l'edat. A Espanya, l'incidència és de 175 casos per 100.000 habitants/any i té una prevalença de 4-8% en les persones majors de 65 anys.

Les malalties vasculars cerebrals es poden dividir en dos grans subgrups, els accidents isquèmics i els hemorràgics [9]. Els accidents isquèmics són deguts a l'oclusió aguda dels vasos cerebrals a causa d'un trombo o d'un èmbol provinent d'un altre territori vascular. Els accidents hemorràgics són secundaris a un vessament de sang intracraneal. La taula 1 mostra la classificació dels AVC.

- (1) **AVC isquèmic:** Oclusió aguda de vasos cerebrals com a conseqüència d'un trombo o d'un èmbol. Al 85% dels pacients que se'ls diagnostica una malaltia vascular cerebral es tracta d'un AVC isquèmic.
- (a) *Global:* Lesions cerebrals difuses que apareixen en situacions d'extrema gravetat com en una parada cardiorrespiratòria.
 - (b) *Focal:*
 - (i) *Accident isquèmic transitori (AIT):* episodis de disminució/alteració neurològica focal de breu duració amb recuperació total del pacient en menys de 24h. Solen ser recurrents.
 - (ii) *Efecte neurològic isquèmic reversible (ENIR):* dèficit neurològic de més de 24h amb recuperació completa en menys de 4 setmanes.
 - (iii) *AVC progressiu:* inici brusc de la simptomatologia i empitjorament clínic progressiu a les primeres hores. No hi ha recuperació posterior.
- (2) **AVC hemorràgic:**
- (a) *Hemorràgia cerebral intraparenquimatososa per hipertensió arterial (HTA):* Col·lecció hemàtica dins el parènquima encefàlic produït per una ruptura vascular amb o sense comunicació amb els espais subaracnoideus o amb el sistema ventricular. La localització més freqüent és a nivell dels ganglis de la base. Tot i que, es pot presentar a qualsevol nivell i al tronc encefàlic.
 - (b) *Hemorràgia subaracnoidea:* Extravasació de sang en l'espai subaracnoide directament.

TAULA 1. Classificació dels ACV.

La taula 2 mostra l'origen més freqüent dels AVC.

Causa	Explicació
Cardioembòlic	Causat per una malaltia cardíaca que allibera èmbols que viatgen a través del sistema circulatori fins a ocluir un vas cerebral.
Arterioscleròtic de gran vas	Engruiximent progressiu de la paret d'una artèria com a resulta de l'acumulació de material gras com el colesterol.
Arterioscleròtic de petit vas o infart llacunar	Malalties sistèmiques com la hipertensió arterial o la diabetis mellitus, deterioren els vasos de petit calibre produint oclusions com a troballa casual.
Causa incerta o mixta.	

TAULA 2. Origen més freqüent dels AVC.

Hi ha nombrosos estudis epidemiològics que han establert els factors de risc de patir un AVC isquèmic. Es mostren a la taula 3.

Modificables	Potencialment modificables	No modificables
Hipertensió arterial	Diabetis Mellitus	Edat avançada
Fumar	Hiperhomocisteïnèmia	Ser home
Malaltia cardíaca	Hipertròfia ventricle esquerre	Antecedents familiars
Fibril·lació auricular		Localització geogràfica
Endocarditis infecciosa		Ètnia
Estenosi mitral		Accident Isquèmic Transitori
Infart agut de miocardi recent		Postmenopausa
Anèmia drepanocítica		
Estenosi carotídia asimptomàtica		
Obesitat		
Sedentarisme		

TAULA 3. Factors de risc dels AVC.

Un pacient amb un AVC pot presentar un o més dels símptomes presentats a la taula 4.

El diagnòstic dels AVC es fa a través de la clínica i es confirma amb les proves d'imatge. El temps entre l'inici del quadre i el moment del diagnòstic és essencial per al pronòstic d'aquests pacients. Les proves d'imatge que es fan són: Tomografia computeritzada (TC), Resonància magnètica nuclear (RMN) i altres.

Síntoma	Explicació
Hemiplegia/hemiparèsia	Paràlisi o parèsia d'una meitat del cos.
Amaurosi fugaç homolateral	Episodi de ceguera monocular transitòria, de curta durada i amb recuperació espontània completa
Afàsia	Afectació de la capacitat de produir i/o comprendre el llenguatge.
Alèxia	Pèrdua de la capacitat per llegir.
Atàxia	descordinació en el moviment de les parts del cos.
Allucinacions visuals	
Trastorns transitoris de la memòria	
Asimptomàtica	

TAULA 4. Símpotmes de l'AVC.

El pronòstic dels pacients amb un AVC és poc favorable i el 25% dels pacients patirà un empitjorament clínic precoç. La mortalitat durant el primer mes oscil·la entre el 8% i el 20% dels pacients mentre que la mortalitat als 5 anys s'eleva al 40-60%. Per millorar el pronòstic és vital reduir el temps a l'iniciar el tractament. A part, els pacient que ingressen en una unitat d'ictus tenen un 18-23% menys de mortalitat.

Hi ha diversos tractaments que es practiquen en els pacients amb un AVC. Aquests es poden agrupar en 5 grans grups presentats a la taula 5.

Tractament	Explicació
Mesures generals. Fibrinòlisi (rt-PA)	Si es practica durant les primeres 3 hores, es redueix la mortalitat i les seqüeles. Però té un risc molt elevat que produeixi una hemorràgia secundària.
Antitrombòtics	Prevé la formació i progressió dels trombus i la recurrència primerenca
Fàrmacs neuroprotectors	Prevé la transformació de l'àrea de penombra isquèmica en infart.
	No s'ha provat la seva eficàcia.
Cirurgia	Es practiquen endoarteriectomies per extraure les plaques d'ateroma.

TAULA 5. Tractament dels pacients amb un AVC.

El mitjà més eficaç per reduir les conseqüències d'un AVC és la prevenció. La prevenció primària és evitar el primer AVC controlant els factors de risc. La prevenció secundària és donar medicaments antiagregants com l'aspirina o el clopidogrel juntament amb anticoagulants (acenocumarol) de per vida als pacients amb un AVC previ.

2. Objectius i Hipòtesis

L'objectiu principal és estimar els efectes de la citicolina en la temperatura corporal. El secundari és avaluar la capacitat pronòstica de la temperatura basal. A nivell personal la meua motivació docent és estudiar quina potser la millor solució i aplicar les tècniques estadístiques més adequades per a respondre un problema real.

Les hipòtesis que es comprovaran són:

- Hipòtesis 1: Els pacients tractats van tenir una disminució de la temperatura corporal més accentuada que els pacients controls.
- Hipòtesis 2: Els pacients amb una temperatura basal elevada van tenir una mortalitat superior als pacients amb una temperatura baixa.

- Hipòtesis 3: Els pacients amb una temperatura basal elevada van acabar l'estudi amb unes seqüeles més severes que els pacients amb temperatures baixes.

Capítol 2

Materials i Mètodes

Aquest apartat té dues parts diferenciades. La primera explica el disseny de l'assaig clínic; la segona, explica els mètodes estadístics.

En la primera part, presentem l'assaig clínic sobre la efectivitat de la citicolina amb pacients que han patit un AVC isquèmic. Abans d'iniciar l'estudi, hem de comprovar que les dades de l'assaig tenen suficient potència per a respondre les tres hipòtesis. A més, mostrarem els diferents patrons de visites per estudiar com les pèrdues poden esbiaixar els resultats.

La segona part expliquem els mètodes lineals mixtos, la regressió logística i la regressió logística ordinal, emprats per a respondre cada una de les tres hipòtesis.

1. Població d'estudi

Les dades provenen d'un estudi fet en 118 centres dels Estats Units i Canadà per a comprovar la seguretat i l'eficàcia de la citicolina en pacients amb un ACV isquèmic. El període d'estudi comprèn des de l'agost del 1998 fins al novembre del 1999. Els criteris d'inclusió més importants eren que el pacient havia de tenir una mínima gravetat inicial i que el tractament comences durant les primeres 24 hores des de l'ACV. La gravetat es va mesurar amb una escala d'alteració, la **Neurologic Institute Health Stroke Scale** (NIHSS), i havien de tenir com a mínim 8 punts.

El disseny de l'estudi va ser el d'un assaig clínic amb emmascarament del pacient i del personal mèdic. L'objectiu principal era determinar l'efecte recuperador de la citicolina amb 6 setmanes de tractament i 6 setmanes més de seguiment. Es van recollir les dades en les següents visites mèdiques: la primera o basal; la segona, després d'acabar la primera setmana de tractament; la tercera, després de la segona setmana de tractament; la quarta; durant la quarta setmana; la cinquena, al completar el tractament; i la última, sis setmanes després d'acabar el tractament.

Al final de l'estudi, un 52% dels pacients tractats amb citicolina van millorar set punts. La proporció en els tractats de placebo va ser d'un 51%. No es van trobar diferències significatives entre els dos grups. Per tant, no es va poder concloure que la citicolina fos un medicament eficaç per als AVC.

1.1. Flux de malalts. L'assaig clínic estava dissenyat per tenir 6 visites al llarg del temps d'estudi de cada pacient. Per tant, l'estudi hauria de tenir 5394 visites agrupades en 899 pacients, però només s'ha recollit les dades de 4089 visites. Existeixen diferents raons per no haver recollit les visites faltants com la mort del pacient, la sortida del pacient de l'estudi, no venir a la visita, etc.

La figura 1 presenta el diagrama de flux dels pacients. Les dues branques de l'estudi tenen 446 (49,6% del total) i 453 (50,4% del total) pacients cada una. Així, l'estudi està balancejat; però només té 289 pacients (32,1% del total) amb el seguiment complet. Dividits en 138 (30,9% dels controls) i 151 (33,3% dels tractats) en cada tractament. El test d'igualtat de proporcions té un p -valor de 0,49. Així, no hi ha prou evidència per a rebutjar que els dos tractaments tenen la mateixa proporció de pacients amb el seguiment completat. A més, tenim 150 pacients (16,7% del total) que van morir al llarg de l'estudi. Repartits en 74 (16,6% dels controls) i 76 (16,8% dels tractats) en els dos tractaments. Tampoc, hi han diferències significatives entre els dos grups (p -valor $\cong 1$).

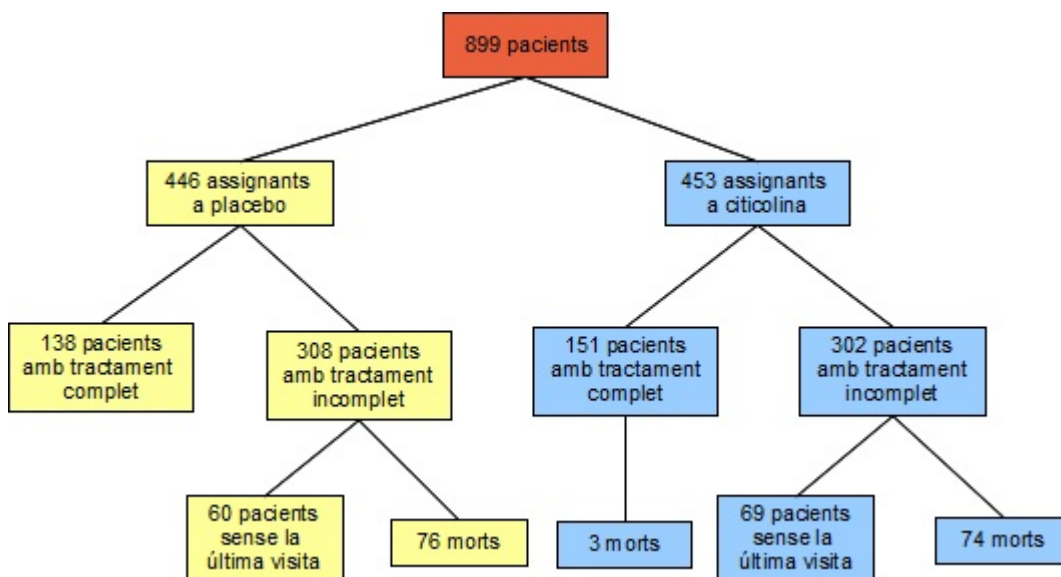


FIGURA 1. Flux de malalts de l'assaig clínic.

La distribució de les pèrdues està representada en la figura 2. En aquesta figura, cada línia representa un dels patrons de visites. La primera línia correspon als pacients que han vingut a totes les visites. Les sis línies següents representen els patrons dels pacients que han faltat només a una visita. Els pacients amb com a molt una sola visita faltant representen el 67,9% (610/899) de tots els pacients. Les visites faltants es concentren en la part final del període d'estudi de cada pacient. Per exemple, els pacients que no han vingut com a mínim a la seva última visita representen un 30,5% (274/899) de tots els pacients.

Al tenir visites mancants, l'estudi ha perdut informació. Una de les possibles conseqüències és que la diferència detectable haurà de ser més gran, ja que augmentarà

Nombre de pacients	Visita						Nombre de pacients	Visita					
	basal	Visita 1	Visita 2	Visita 3	Visita 6	Visita 12		basal	Visita 1	Visita 2	Visita 3	Visita 6	Visita 12
289	-----	-----	-----	-----	-----	-----	1	-	-----	-	-	-----	-----
13	-	-----	-----	-----	-----	-----	2	-----	-	-----	-	-	-----
211	-----	-	-----	-----	-----	-----	1	-----	-----	-	-	-	-----
22	-----	-----	-	-----	-----	-----	1	-	-	-----	-----	-----	-
29	-----	-----	-----	-----	-	-----	1	-	-----	-	-----	-----	-
14	-----	-----	-----	-----	-	-----	5	-----	-	-----	-	-----	-
32	-----	-----	-----	-----	-----	-	3	-	-----	-----	-----	-	-
6	-	-	-----	-----	-----	-----	11	-----	-	-----	-----	-----	-
3	-	-----	-	-----	-----	-----	12	-----	-----	-	-----	-	-
2	-----	-	-	-----	-----	-----	18	-----	-----	-----	-	-	-
9	-----	-	-----	-	-----	-----	1	-----	-	-	-	-----	-
2	-	-----	-----	-----	-	-----	1	-	-	-----	-----	-----	-
7	-----	-	-----	-----	-	-----	1	-	-----	-	-----	-	-
3	-----	-----	-	-----	-	-----	2	-----	-	-	-----	-	-
4	-----	-----	-----	-	-	-----	3	-	-----	-----	-	-	-
2	-	-----	-----	-----	-----	-	22	-----	-	-----	-	-	-
24	-----	-	-----	-----	-----	-	34	-----	-----	-	-	-	-
8	-----	-----	-	-----	-----	-	1	-	-	-	-----	-	-
6	-----	-----	-----	-	-----	-	1	-	-	-----	-	-	-
20	-----	-----	-----	-----	-	-	3	-	-----	-	-	-	-
2	-	-	-----	-	-----	-----	62	-----	-	-	-	-	-

FIGURA 2. Distribució de les visites perdudes. El símbol $-|$ significa que el pacient va venir a la visita i $-$ que va faltar a la visita.

l'error estàndard de les mesures que es comparin. Un altre conseqüència és que les pèrdues poden ser no aleatòries provocant possiblement un biaix per una banda i, per l'altra, buidant de significat les mesures de l'error d'estimació, que ja no és únicament aleatori.

1.2. Potència de les dades. L'objectiu de l'assaig clínic era la comparació de la proporció de pacients amb una millora de set punt en l'escala NIHSS en dotze setmanes en els dos grups. S'esperava que un 51% dels pacients tractats amb citicolina millorarien set punts, enfront d'un 41% en el tractats amb placebo. Per tant, el tamany mostral necessari és de 900 pacients en els dos grups si el test és de dues cues amb un error tipus I de 0,05 i una potència del 0,85.

L'hipòtesis principal del present estudi és la comparació de les temperatures mitjanes, no les proporcions de morts. En aquest cas, necessitem un tamany mostral per cada mostra de

$$n = \frac{2 \cdot \sigma^2 \cdot (\Phi(1 - \alpha/2) + \Phi(potencia))^2}{\Delta^2}$$

on σ és la variància comuna a cada mostra, $\Phi(x)$ és el quantil de la $P(Y < x)$ segons una $N(0, 1)$ i Δ és la diferència esperada.

Podem saber quina diferència mínima es pot detectar segons el tamany mostral, la variància i la potència. També es pot saber quina potència es tindrà segons el tamany mostral, la variància i la diferència mínima esperada. La taula 1 presenta les taules per a la diferència mínima segons el tamany mostral ($N = 900$), una potència del 80% i diverses desviacions estàndard de la temperatura corporal. També inclou l'estudi de la potència segons el tamany mostral, la diferència mínima esperada i diverses desviacions estàndards.

Diferència mínima detectables amb una potència de 0,8		Potència segons la desviació estàndard i la diferència mínima observada				
Desviació estàndard	Diferència mínima	Desviació estàndard	Diferència mínima			
			0,100	0,050	0,025	0,010
0,10	0,0187	0,10	1,00	1,00	0,96	0,32
0,25	0,0467	0,25	1,00	0,85	0,32	0,09
0,50	0,0933	0,50	0,96	0,32	0,11	0,05
0,75	0,1401	0,75	0,51	0,17	0,07	0,04
1,00	0,1868	1,00	0,32	0,11	0,06	0,03

TAULA 1. La taula esquerra, hi ha la diferència mínima segons un tamany mostral total de 900 pacients, la desviació estàndard i una potència del 80%. La taula dreta presenta la potència segons el tamany mostral, una desviació estàndard i la diferència mínima esperada.

Amb el tamany mostral i una potència del 80%, no són detectables diferències inferiors a $0,0187^{\circ}C$ amb una desviació estàndard de la temperatura de $0,1^{\circ}C$. Per tant, les diferències entre les temperatures dels dos grups hauran de ser de l'ordre de les dècimes per a poder-les detectar, ja que segurament tindrem una desviació estàndard superior.

Si ens fixem en la potència segons la desviació estàndard i la diferència mínima, la majoria de combinacions donen potències baixes. Les potències més elevades es concentren en la desviació estàndard de $0,1^{\circ}C$ o si les diferències mínimes són superiors a $0,1^{\circ}C$. Per tant, si la diferència mínima és inferior o la desviació estàndard és superior, és possible que la potència de les dades no sigui suficient per a rebutjar la hipòtesis d'igualtat de les mitjanes.

Hem de recordar; però que l'anàlisi serà més sofisticat que una comparació de dues mitjanes amb un t -test i que l'amplada dels intervals de confiança al 95% de l'efecte estimat serà una estimació complementària de la informació aportada per aquest estudi.

2. Mètodes estadístics

Per a respondre els tres objectius, necessitem utilitzar diferents mètodes estadístics. Per modelar la temperatura necessitem un model de regressió lineal amb resposta en els reals. En canvi, per a estudiar la mort o les seqüeles necessitem models de regressió amb resposta binària o ordinal, respectivament. A més, en els tests i intervals de confiança s'ha fixat un error tipus I de 0,05.

2.1. Models lineals mixtos. El model lineal és un model clàssic i dels més utilitzats. Té tres grans supòsits: el primer és que la resposta és una variable real; el segon, que la relació entre la variable resposta i els paràmetres és lineal; i el tercer, que els residus es distribueixen com una multinormal centrada en el zero i amb

una variància constant i covariàncies nul·les. Els models lineals mixtos relaxen la tercera condició.

Els models lineals mixtos ([21], [22], [7]) són models en els qual tant la part fixa com la aleatòria tenen un comportament lineal entre els paràmetres i la resposta. Utilitzarem la modelització de Laird i Ware [14] explicada a [19].

2.1.1. *Formulació.* El model lineal mixt té dues parts diferenciades. La primera part és un model lineal clàssic. La segona modela els residus del model anterior de manera lineal. El model per a l' i -èssim grup s'expressa com

$$(1) \quad \begin{aligned} y_i &= X_i\beta + Z_ib_i + \epsilon_i, \quad i = 1, \dots, M \\ b_i &\sim N(0, \Psi), \epsilon_i \sim N(0, \sigma^2 I) \end{aligned}$$

on la resposta del grup i , y_i , és un vector de llargada n_i , β és el vector d'efectes fix de llargada p , b_i és el vector d'efectes aleatoris de llargada q , X_i (dim $n_i \times p$) i Z_i (dim $n_i \times q$) són les matrius dels efectes fixos i aleatoris i ϵ_i (dim n_i) és el vector d'errors dins el grup. Suposarem que la distribució dels errors dins del grup segueix una distribució normal multivariant amb esperança 0 i variància $\sigma^2 I$. La condició sobre la variància dels errors dintre del grup es pot relaxar com veurem en el final d'aquesta secció. A més, les components aleatòries, b_i , i l'error dintre del grup, ϵ_i , són independents en cada grup.

Suposem que els efectes aleatoris, b_i , es distribueixen com una normal multivariant amb esperança 0 i una matriu de covariàncies Ψ . La matriu de covariàncies caracteritza completament els efectes aleatoris. A més, la matriu de covariàncies és millor expressar-la respecte del *factor de precisió relatiu*, Δ , que satisfà

$$\frac{\Psi^{-1}}{1/\sigma^2} = \Delta^T \Delta.$$

2.1.2. *Estimació.* Els dos mètodes més utilitzats per estimar els paràmetres d'un model lineal mixt són el mètode de màxima versemblança (MV) i el de màxima versemblança restringida (MVR).

La funció de versemblança del model és la densitat de probabilitat de les dades donat els paràmetres. Així,

$$(2) \quad L(\beta, \theta, \sigma^2 | y) = p(y | \beta, \theta, \sigma^2)$$

on L és la versemblança, p és la densitat de la probabilitat, y és el N -vector resposta, $N = \sum_i n_i$, i θ són els paràmetres referents als efectes aleatoris. Per calcular la versemblança utilitzarem la densitat marginal, $p(y | \beta, \theta, \sigma^2)$.

El model lineal mixt es pot reexpressar com

$$y_i = X_i\beta + \epsilon'_i, \quad i = 1, \dots, M$$

on $\epsilon'_i = Z_ib_i + \epsilon_i$. L'error ϵ'_i és la suma de dos vectors normals multivariants independents que es distribueix com una normal multivariant amb mitjana 0 i matriu de covariàncies $\sigma^2 E$ on $E = I + Z_i\Psi Z_i^T / \sigma^2$. Per tant, y_i té una mitjana $X_i\beta$ i una matriu de covariàncies $\sigma^2 E$. Per tant, la versemblança del model és

$$(3) \quad p(y_i | \beta, \theta, \sigma^2) = (2\pi\sigma^2)^{-n_i/2} \exp\left(\frac{(y_i - X_i\beta)^T E_i^{-1} (y_i - X_i\beta)}{-2\sigma^2}\right) |E_i|^{-1/2}.$$

Donat un θ , els valors màxim versemblants de β i σ^2 són

$$(4) \quad \beta(\theta) = \left(\sum_{i=1}^M X_i^T E_i^{-1} X_i \right)^{-1} \sum_{i=1}^M X_i^T E_i^{-1} y_i,$$

$$(5) \quad \sigma^2(\theta) = \frac{\sum_{i=1}^M (y_i - X_i \beta(\theta))^T E_i^{-1} (y_i - X_i \beta(\theta))}{N}.$$

Però no utilitzarem aquestes expressions per obtenir els paràmetres, ja que són computacionalment costoses d'avaluar. La versemblança es calcula a través de l'aproximació utilitzant la pseudo-data. Definim la pseudo-data com

$$(6) \quad \tilde{y}_i = \begin{bmatrix} y_i \\ 0 \end{bmatrix}, \quad \tilde{X}_i = \begin{bmatrix} X_i \\ 0 \end{bmatrix}, \quad \tilde{Z}_i = \begin{bmatrix} Z_i \\ 0 \end{bmatrix} \quad \text{i} \quad \tilde{b}_i = (\tilde{Z}_i^T \tilde{Z}_i)^{-1} \tilde{Z}_i^T (\tilde{y}_i - \tilde{X}_i \beta).$$

Aleshores, la versemblança es pot expressar com

$$(7) \quad L(\beta, \theta, \sigma^2 | y) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp \left(-\frac{\sum_{i=1}^M |\tilde{y}_i - \tilde{X}_i \beta - \tilde{Z}_i \tilde{b}_i|^2}{2\sigma^2} \right) \frac{\prod_{i=1}^M \|\Delta\|}{\sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}.$$

Per tant, utilitzant les expressions de $\beta(\theta)$ i $\sigma^2(\theta)$ tenim la versemblança “profiled” que s'expressa com

$$(8) \quad L(\theta) = L(\beta(\theta), \theta, \sigma^2(\theta)) = \frac{\exp(-N/2) \prod_{i=1}^M |\Delta|}{(2\pi\sigma^2(\theta))^{N/2} \sqrt{|\tilde{Z}_i^T \tilde{Z}_i|}}.$$

Per avaluar la versemblança “profiled” no necessitem conèixer ni els efectes aleatoris ni $\beta(\theta)$. Només es necessita conèixer la norma dels residus del problema de mínims quadrats que s'obté utilitzant una descomposició QR.

L'estimació màxim versemblant té el problema de sota-estima la variància residual. Per solucionar aquest problema es pot utilitzar l'estimació màxim versemblant restringida. Es defineix la versemblança restringida com

$$(9) \quad L_R(\theta, \sigma^2 | y) = \int L(\beta, \theta, \sigma^2 | y) d\beta.$$

El context Bayesià assumeix que la distribució prèvia és localment uniforme per als efectes fixos.

Per optimitzar el model s'utilitza el logaritme de la versemblança “profiled” o de la versemblança restringida “profiled”. Els mètodes per optimitzar són l'algoritme *EM* i el mètode de Newton-Raphson.

L'algoritme *EM* té dos passos per cada iteració. El primer pas, *E*, és utilitzar els paràmetres $\theta^{(w)}$ per avaluar la distribució $b|y$ i tenir l'esperança del logaritme de la versemblança per θ . El segon pas, *M*, consisteix en maximitzar l'esperança respecte θ per produir una nova $\theta^{(w+1)}$. Aquest algoritme és molt utilitzat per a fer estimacions màxim versemblants quan en les dades hi han buits.

El mètode de Newton-Raphson utilitza una expansió de Taylor de primer ordre al voltant de $\theta^{(w)}$ per produir una nova estimació, $\theta^{(w+1)}$. En cada iteració de l'algoritme és necessari calcular el logaritme de la versemblança, el seu gradient i la seva Hessiana. És un algoritme computacionalment costós, però té una convergència quadràtica. Normalment, s'utilitzen els algoritmes quasi-Newton. Aquests a cada iteració actualitzen una pseudo-Hessiana, en comptes de calcular la Hessiana cada vegada.

L'algoritme *EM* arriba ràpidament a la regió òptima dels paràmetres, però en aquesta regió li costa convergir als paràmetres òptims. En canvi, l'algoritme de Newton-Raphson és computacionalment costós i inestable lluny de la regió òptima dels paràmetres. Per tant, utilitzarem els mètodes híbrids els quals inicialitzen l'algoritme de Newton-Raphson amb una iteració de l'algoritme *EM*.

2.1.3. *Inferència.* La inferència dels paràmetres d'un model lineal mixt es fa a través d'aproximacions a una distribució normal, [18].

Els estimadors màxim versemblants es distribueixen asimptòticament com

$$(10) \quad \begin{pmatrix} \tilde{\beta} \\ \tilde{\theta} \\ \tilde{\sigma}^2 \end{pmatrix} \sim N \left[\begin{pmatrix} \beta \\ \theta \\ \sigma^2 \end{pmatrix}, \begin{pmatrix} \sigma^2(X^T EX)^{-1} & 0 \\ 0 & I_{\theta, \sigma^2}^{-1} \end{pmatrix} \right]$$

amb la matriu d'informació de Fisher expressada com $I_{\theta, \sigma^2} = -E \left[\frac{\delta^2 L(\beta, \theta, \sigma^2 | y)}{\delta \theta \delta \sigma^2} \right]$. Per a distribuir-se asimptòticament s'han de satisfer unes condicions que s'expliquen a [18].

2.1.4. *Contrast d'hipòtesis i intervals de confiança.* Un cop s'ha estimat un model volem saber la precisió de l'estimació i si els paràmetres són significatius. Per establir la significació dels paràmetres utilitzarem el test de raó de versemblances i de Wald. La precisió dels paràmetres l'establirem amb els intervals de confiança.

Test de raó de versemblances. El test de la raó de versemblances serveix per comparar models estimats amb màxima versemblança i l'especificació dels efectes fixos ha de ser comuna a tots els models. Els models que comparem han d'estar anidats. És a dir, un dels dos models ha de representar un cas especial de l'altre. Utilitzarem aquest test per estimar la significació dels efectes aleatoris del model.

Si L_2 és la versemblança del model més general i L_1 és la versemblança del model restringit, es satisfà que $L_2 > L_1$. L'estadístic del test de raó de versemblances (*LRT*) és

$$LRT = 2 \cdot \log(L_2/L_1) = 2 [\log(L_2) - \log(L_1)]$$

i serà positiu. Si k_1 és el nombre de paràmetres del model i , la distribució de l'estadístic *LRT* sota la hipòtesis nul·la serà $\chi_{k_2 - k_1}^2$.

Per comparar models també es poden utilitzar els criteris d'informació d'Akaike (*AIC*) o el Bayesià (*BIC*). Aquests criteris corregeixen la versemblança en funció del nombre de paràmetres del model. Així,

$$\begin{aligned} AIC &= -2L(\tilde{\theta}|y) + 2n_{par}, \\ BIC &= -2L(\tilde{\theta}|y) + n_{par} \log(N). \end{aligned}$$

on n_{par} és el nombre de paràmetres de cada model i N el número d'observacions. Quan es comparen dos models, el millor és el que té el *AIC* o *BIC* mínim.

Test de Wald. Per estimar la significació dels efectes fixos, utilitzarem el test de Wald. Aquest test només es pot aplicar a estimadors màxim versemblants. Les hipòtesis del test són

$$\begin{aligned} H_0 &: \beta = \beta_0 \\ H_1 &: \beta \neq \beta_0 \end{aligned}$$

Sota la hipòtesis nul·la, l'estimador màxim versemblant de β es distribueix com una distribució normal, $\tilde{\beta} \sim N(\beta_0, I(\beta_0)^{-1})$. Per tant, es defineix l'estadístic com

$$(\beta - \beta_0)^T I(\beta_0)^{-1} (\beta - \beta_0)$$

i sota la hipòtesis nul·la es distribueix com una distribució χ_p^2 on p és la dimensió de β .

Intervals de confiança Els intervals de confiança es calculen sota el supòsit de la distribució asimptòtica. Per tant, els intervals de confiança per als efectes fixos són

$$\tilde{\beta}_j \pm t_{df_j}(1 - \alpha/2) \tilde{\sigma}_R \sqrt{[(X^T EX)^{-1}]_{jj}}$$

on $t_{df_j}(1 - \alpha/2)$ és el quartil $1 - \alpha/2$ d'una distribució t amb df_j graus de llibertat, df_j és la dimensió de β_j , $\tilde{\sigma}_R$ és l'estimació màxim versemblant restringida de σ i $[(X^T EX)^{-1}]_{jj}$ és l'element de la fila j i la columna j de la matriu $(X^T EX)^{-1}$.

L'interval de confiança per a la variància intra-grup també s'extreu de la distribució asimptòtica i s'expressa com

$$\tilde{\sigma} \exp \left(\pm z(1 - \alpha/2) \sqrt{[I^{-1}]_{\sigma\sigma}} \right)$$

on $[I^{-1}]_{\sigma\sigma}$ representa l'últim element de la diagonal de l'inversa de la matriu d'informació i $z(1 - \alpha/2)$ és el quartil $1 - \alpha/2$ d'una normal estàndard.

Finalment, els intervals de confiança per als elements de la diagonal de la matriu de covariàncies dels efectes aleatoris són

$$\exp(\tilde{\eta}_j \pm z(1 - \alpha/2) \sqrt{[I^{-1}]_{jj}})$$

i pels de fora de la diagonal són

$$\left[\frac{\exp(\tilde{\eta}_j - z(1 - \alpha/2) \sqrt{[I^{-1}]_{jj}}) - 1}{\exp(\tilde{\eta}_j - z(1 - \alpha/2) \sqrt{[I^{-1}]_{jj}}) + 1}, \frac{\exp(\tilde{\eta}_j + z(1 - \alpha/2) \sqrt{[I^{-1}]_{jj}}) - 1}{\exp(\tilde{\eta}_j + z(1 - \alpha/2) \sqrt{[I^{-1}]_{jj}}) + 1} \right].$$

2.1.5. *Extensions del model lineal.* Quan els residus intra-grup són heterocedàstics o estan correlats, el model lineal mixt no és capaç de captar aquesta font de variació. El model lineal mixt es pot flexibilitzar per a poder modelar aquesta font de variació.

Formulació del model lineal mixt general El model lineal mixt suposa que els errors intra-grup es distribueixen com una normal amb mitjana 0 i variància un múltiple de la identitat. Flexibilitzar el model significa relaxar l'assumpció d'homocedàsticitat dels residus intra-grup. Per tant, el model lineal mixt ampliat serà

$$(11) \quad \begin{aligned} y_i &= X_i\beta + Z_ib_i + \epsilon_i, \quad i = 1, \dots, M \\ b_i &\sim N(0, \Psi), \quad \epsilon_i \sim N(0, \sigma^2\Lambda_i) \end{aligned}$$

on Λ_i és una matriu semi-definida positiva parametritzada pel conjunt de paràmetres λ . Com en el model lineal mixt, els efectes aleatoris i els errors intra-grup són independents.

La variància de les observacions s'expressa com

$$V(y_i) = \Sigma_i = \sigma^2(Z_iDZ_i^T + \Lambda_i).$$

La component $Z_iDZ_i^T$ correspon als efectes aleatoris i Λ_i a la variància intra-grup. Fins ara, ens hem centrat en els efectes aleatoris i s'ha suposat que $\Lambda_i = I$.

La matriu Λ_i es pot descomposar en la multiplicació de matrius més simples,

$$\Lambda_i = V_iC_iV_i.$$

La matriu V_i és diagonal i C_i és una matriu semi-definida positiva amb els elements de la diagonal iguals a 1. Es pot comprovar que

$$Var(\epsilon_{ij}) = \sigma^2 [V_i]_{jj}^2, \quad Cor(\epsilon_{ij}, \epsilon_{ik}) = [C_i]_{jk}.$$

Aquesta descomposició en una matriu de variàncies i una altra de correlacions és convenient per desenvolupar el marc teòric. Ens permet modelar les dues estructures per separat.

Funcions per a modelar la heterocedasticitat Les funcions de variància ens serviran per a modelar la heterocedasticitat del model. Definim la funció de variància com

$$Var(\epsilon_{ij}|b_i) = \sigma^2 g^2(\mu_{ij}, v_{ij}, \delta), \quad i = 1, \dots, M, \quad j = 1, \dots, n_i.$$

on $\mu_{ij} = E[y_{ij}|b_i]$, v_{ij} és el vector de les variàncies de les covariables, δ és el paràmetre de la variància i $g(\cdot)$ és la funció de variància. Se suposa que la funció de variància és contínua sobre δ .

Les diferents possibles funcions de variància són:

- Variància fixa: $Var(\epsilon_{ij}) = \sigma^2 v_{ij}$ on v_{ij} és la variància de la covariable i $g(v_{ij}, \delta) = \sqrt{v_{ij}}$.
- Diferents variàncies per estrat: $Var(\epsilon_{ij}) = \sigma^2 \delta_{S_{ij}}^2$ on S_{ij} és l'estrat de l'observació y_{ij} i $g(S_{ij}, \delta) = \delta_{S_{ij}}$.
- Potència d'una covariable: $Var(\epsilon_{ij}) = \sigma^2 |v_{ij}|^{2\delta}$ on $g(v_{ij}, \delta) = |v_{ij}|^\delta$.
- Exponencial d'una covariable: $Var(\epsilon_{ij}) = \sigma^2 \exp(2\delta v_{ij})$ on $g(v_{ij}, \delta) = \exp(\delta v_{ij})$.
- Potència d'una covariable més una constant: $Var(\epsilon_{ij}) = \sigma^2 (\delta_1 + |v_{ij}|^{\delta_2})^2$ on $g(v_{ij}, \delta) = \delta_1 + |v_{ij}|^{\delta_2}$ i $\delta_1 > 0$.
- Combinació de diverses funcions: és qualsevol combinació de les funcions anteriors.

Estructura de correlació La estructura de correlacions entre les observacions intra-grup es pot expressar com

$$Cor(\epsilon_{ij}, \epsilon_{ik}) = h[d(p_{ij}, p_{ik}), \rho] \quad i = 1, \dots, M \text{ i } j, k = 1, \dots, n_i$$

on ρ és el vector de paràmetres de les correlacions, $h(\cdot)$ és una funció de correlació que pren valors entre $[-1, 1]$, contínua sobre ρ i tal que $h(0, \rho) = 1$. El vector p_{ij} representa la posició de ϵ_{ij} i es pot calcular la distància entre dues posicions amb $d(\cdot, \cdot)$.

Restringirem les correlacions intra-grup a una correlació de sèrie temporal. Així, suposem que els errors intra-grup es distribueixen com una sèrie temporal. La correlació serial es pot definir com

$$Cor(\epsilon_{ij}, \epsilon_{ik}) = h(|p_{ij} - p_{ik}|, \rho) \quad i = 1, \dots, M \text{ i } j, k = 1, \dots, n_i.$$

En el context de les sèries temporals la funció s'anomena funció d'autocorrelacions. Les funcions d'autocorrelació són:

- *Compound Symmetry*: és la funció més simple i suposa una correlació constant dintre de cada grup. És a dir, $Cor(\epsilon_{ij}, \epsilon_{ik}) = \rho \forall i \neq k$, $h(j - k, \rho) = \rho$ i ρ s'anomena el coeficient de correlació intra-classe.
- *General*: cada correlació està definida per un paràmetre diferent. Així, $h(j - k, \rho) = \rho_{j-k}$. Aquesta funció sobre-parametriza l'estructura de correlació.
- *ARMA*: es pot modelar la sèrie dels errors intra-grup com un model autoregressiu de mitjanes mòbils. Tot i que es pot definir qualsevol model *ARMA*, només explicarem els autoregressiu d'ordre 1 (*AR(1)*), els més freqüents. Els models *AR(1)* tenen una funció d'autocorrelació senzilla, $h(i - k, \phi) = \phi^{i-k}$ on ϕ representa la correlació entre $\epsilon_{i,j}$ i $\epsilon_{i,j-1}$. Quan $i - k$ no és un enter, tenim la funció contínua autoregressiva d'ordre 1 (*CAR(1)*). Però en aquest cas, el paràmetre ϕ no pot ser negatiu.

2.1.6. *Software*. Els models lineals mixtos s'han estimat amb el programa "R version 2.11.1". Les rutines específiques estan a la llibreria "nlme version 3,1-96".

2.2. Model de regressió logístic. Quan el model té una resposta dicotòmica, el model més utilitat per a estimar l'efecte de les covariables sobre la resposta és el model de regressió logístic [12].

2.2.1. *Formulació*. Si disposem d'un vector de p variables independents, $x = (x_1, \dots, x_p)$, i una variable dependent dicotòmica, el model de regressió logística s'expressa com

$$(12) \quad \text{logit}(\pi(x)) = \ln \left(\frac{\pi(x)}{1 - \pi(x)} \right) = \alpha + \beta_1 x_1 + \dots + \beta_p x_p + \epsilon(x)$$

on $\pi(x) = P(y = 1|x)$. Si algunes de les p variables és categòrica amb q nivells, s'introduran $q - 1$ variables dicotòmiques per a captar l'efecte de la variable. Així, la variable amb q categories x_i es substitueix per $q - 1$ variables indicadores una per a cada nivell,

$$x_{i,1} = 1(x_i = 2), \dots, x_{i,q-1} = 1(x_i = q).$$

Per tant, el model s'expressa com

$$\text{logit}(\pi(x)) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \alpha + \beta_1 x_1 + \dots + \sum_{j=1}^{q-1} \beta_{i,j} x_{i,j} + \dots + \beta_p x_p + \epsilon(x).$$

2.2.2. *Estimació.* Suposem que tenim una mostra aleatòria de n observacions independents, (x_i, y_i) $i = 1, \dots, n$, i volem estimar els $p+1$ paràmetres, $\beta = (\alpha, \beta_1, \dots, \beta_p)$. El mètode més usual per estimar els paràmetres és el de màxima versemblança. La funció de versemblança per al model de regressió logística és

$$(13) \quad L(\beta|y) = \ln(l(\beta|y)) = \sum_{i=1}^n [y_i \ln(\pi(x_i)) + (1 - y_i) \ln(1 - \pi(x_i))].$$

on $\pi(x) = \frac{\exp(\alpha + \sum \beta_i x_i)}{1 + \exp(\alpha + \sum \beta_i x_i)}$. Quan diferenciem la versemblança respecte els $p + 1$ paràmetres obtenim $p + 1$ equacions,

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \text{ i}$$

$$\sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0 \text{ per } j = 1, \dots, p.$$

La solució de les $p+1$ equacions són els paràmetres que maximitzen la versemblança.

Per obtenir les variàncies i covariàncies dels paràmetres, utilitzarem la matriu d'informació de Fisher. Aquesta matriu s'obté de la Hessiana de la versemblança que s'expressa com

$$I_{jj} = \frac{\partial^2 L(\beta)}{\partial \beta_j^2} = - \sum_{i=1}^n x_{ij}^2 \pi(x_i)(1 - \pi(x_i)) \text{ i}$$

$$I_{jl} = \frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_l} = - \sum_{i=1}^n x_{ij} x_{il} \pi(x_i)(1 - \pi(x_i)) \text{ per } j, l = 1, \dots, p.$$

Així, la matriu de variàncies es defineix com $\text{Var}(\beta) = I(\beta)^{-1}$. Una manera alternativa de definir la matriu d'Informació de Fisher és $\hat{I}(\beta) = X^T V X$ on X conté les dades per a cada individu i V és una matriu diagonal omplerta amb $\pi(x_i)(1 - \pi(x_i))$.

2.2.3. *Inferència.* Quan el model està estimat, hem de comprovar que el model no conté elements superflus o que és suficientment bo. Per conèixer la significància dels elements utilitzarem el test de la deviància o de Wald. També calcularem la precisió dels paràmetres amb els intervals de confiança. Finalment, per comprovar la bondat del model construirem diverses mesures objectives per a poder confirmar que el model és adequat.

Test de la deviància Aquest test ens serveix per contrastar si les covariables són significatives. El test compara el model amb les covariables amb un model sense. Definim la deviància com

$$D(\beta) = -2 \ln(l(\beta|y)).$$

Si volem contrastar si un conjunt de p covariables, $\{\beta_i, \dots, \beta_j\}$, tenen un efecte no significatiu, utilitzarem les deviancies dels dos models. El canvi de deviances ens permet construir l'estadístic del test,

$$G = D(\beta \setminus \{\beta_i, \dots, \beta_j\}) - D(\beta) = -2 \ln \left(\frac{l(\beta \setminus \{\beta_i, \dots, \beta_j\} | y)}{l(\beta | y)} \right).$$

Sota la hipòtesis nul·la, es distribueix com una χ_p^2 .

Test de Wald El test de Wald ens permet contrastar la hipòtesis de significació d'un sol paràmetre. Així, les hipòtesis són

$$\begin{aligned} H_0 &: \beta_i = \beta_0 \\ H_1 &: \beta_i \neq \beta_0. \end{aligned}$$

L'estadístic de Wald és $W(\beta_i) = \frac{(\beta_i - \beta_0)^2}{\text{Var}(\beta_i)}$ que sota la hipòtesis nul·la es distribueix com χ_1^2 .

Intervals de confiança Els intervals de confiança es construeixen gràcies a la distribució asimptòtica dels paràmetres. Així, cada paràmetre asimptòticament segueix una normal amb esperança el paràmetre i variància l'element de la diagonal de la matriu inversa d'informació de Fisher. Els intervals de confiança s'expressen com a

$$\beta_i \pm 1,96 \sqrt{\text{Var}(\beta_i)}.$$

Mesures de bondat Per establir si el model és adequat o no, podem quantificar com de bo és el model amb les mesures de bondat. Aquestes mesures ens informen de si el model es capaç de reproduir les dades reals. Utilitzarem dues mesures de bondat, la deviancia i el test de Hosmer-Lemeshow, i una mesura de la capacitat predictiva, l'àrea sota la corba ROC.

Suposem que el nostre model conté p variables independents, $x^T = (x_1, \dots, x_p)$, i sigui J el nombre de valors diferents observats en x . El nombre d'individus que tenen les variables iguals a $x = x_j$ és m_j per $j = 1, \dots, J$ i $\sum m_j = n$, el nombre d'observacions total. A més, notem per y_j el nombre d'esdeveniments, $y = 1$, en el conjunt $x = x_j$ i $\sum y_j = n_1$, el nombre d'esdeveniments en tota la mostra. Totes les distribucions sobre les mesures de bondat són asimptòtiques.

La deviancia i l'estadístic χ^2 de Pearson es construeixen a partir dels residus del model. Aquests s'expressen com

$$r(y_j, \hat{\pi}_j) = \frac{y_j - m_j \hat{\pi}_j}{\sqrt{m_j \hat{\pi}_j (1 - \hat{\pi}_j)}}$$

i l'estadístic de Pearson es distribueix com una χ_{J-p-1}^2 i es defineix com la suma de quadrats dels residus de cada patró, $X^2 = \sum_{j=1}^J r(y_j, \hat{\pi}_j)^2$.

També es pot construir un estadístic a partir de la deviancia residual com la suma de quadrats de la deviancia residual de cada patró, $D = \sum_{j=1}^J d(y_j, \hat{\pi}_j)^2$. Les $d(y_j, \hat{\pi}_j)$ es defineixen segons la resposta dintre de cada patró. Si $y_j = 0$, es defineix com $d(y_j, \hat{\pi}_j) = -\sqrt{2m_j |\ln(1 - \hat{\pi}_j)|}$. En canvi, si $y_j = m_j$, la deviancia residual

és $d(y_j, \hat{\pi}_j) = \sqrt{2m_j |\ln(\hat{\pi}_j)|}$. L'estadístic de les deviàncies residuals es distribueix asimptòticament igual que l'estadístic de Pearson.

Hosmer i Lemeshow van crear una mesura de la bondat basant-se en el test χ^2 per a taules $g \times 2$. L'estadístic es defineix com

$$\hat{C} = \frac{\sum_{k=1}^g (o_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)}$$

on n'_k és el nombre de subjectes dintre del k -èssim grup, c_k denota el nombre de patrons de les covariables en el k -èssim grup, $o_k = \sum_{j=1}^{c_k} y_j$ és el nombre de respostes entre els c_k patrons de covariables i $\bar{\pi}_k = \frac{\sum_{j=1}^{c_k} m_j \hat{\pi}_j}{n'_k}$ és la mitjana de la probabilitat estimada. Definim els g grups com els g -percentils de la probabilitat estimada. L'estadístic \hat{C} es distribueix asimptòticament com una χ_{g-2}^2 . El p -valor del test ens donarà un valor entre 0 i 1 que mesura la bondat del model. Com més aprop de 1 més bo és el model.

L'àrea sota la corba ROC mesura l'habilitat del model de discriminar entre els subjectes que tenen l'esdeveniment dels que no. La corba ROC ve definida per la sensibilitat i la especificitat. Els punts de la corba ROC es defineixen com $(1 - \text{especificitat}, \text{sensibilitat})$. L'estadístic ROC és l'àrea sota la corba ROC. Segons el valor de l'àrea podem concloure si el model discrimina adequadament. Com a norma general [12] tenim:

- Si $ROC = 0,5$, el model no discrimina.
- Si $0,7 \leq ROC \leq 0,8$, el model discrimina de manera acceptable.
- Si $0,8 \leq ROC \leq 0,9$, el model discrimina de manera notable.
- Si $0,9 \leq ROC$, el model discrimina de manera excel·lent.

2.2.4. *Interpretació dels paràmetres.* Per interpretar l'efecte d'una covariable en un model logístic no ens centrarem en el paràmetre estimat. Ens fixarem en l'odds ratio (OR) de l'efecte. Per a una variable dicotòmica, per a una categoria d'una categòrica o per l'augment d'una unitat d'una quantitativa podem estimar l' OR com

$$\begin{aligned} OR(x_i) &= \frac{P(y = 1|x, x_i = 1)/(1 - P(y = 1|x, x_i = 1))}{P(y = 1|x, x_i = 0)/(1 - P(y = 1|x, x_i = 0))} \\ &= \frac{\left(\frac{\exp(\beta_0 + \beta_i + \sum_{j \neq i} \beta_j x_j)}{1 - \exp(\beta_0 + \beta_i + \sum_{j \neq i} \beta_j x_j)} \right)}{\left(\frac{\exp(\beta_0 + \sum_{j \neq i} \beta_j x_j)}{1 - \exp(\beta_0 + \sum_{j \neq i} \beta_j x_j)} \right)} / \frac{\left(\frac{1}{1 - \exp(\beta_0 + \beta_i + \sum_{j \neq i} \beta_j x_j)} \right)}{\left(\frac{1}{1 - \exp(\beta_0 + \sum_{j \neq i} \beta_j x_j)} \right)} = \exp(\beta_i) \end{aligned}$$

amb un interval de confiança $CI(OR, 95\%) = \exp(\beta_i \pm 1,96\sqrt{Var(\beta_i)})$. Construït suposant que el paràmetre es distribueix asimptòticament com una normal. Per a una variable contínua l' OR depèn de les c unitats que se separa de la variable x_j . Així, $OR(x_j + c, x_j) = \exp(c\beta_j)$ amb un interval de confiança de $CI(OR(x_j + c, x_j), 95\%) = \exp(\beta_j \pm 1,96c\sqrt{Var(\beta_j)})$.

2.3. Model de regressió logístic ordinal. Per estudiar les seqüeles d'una AVC s'utilitza l'escala mRS . Així, la resposta del model no és ni una variable real ni una

dicotòmica, és una ordinal. Necessitem un model capaç de modelar una resposta ordinal com el model de regressió logística ordinal [12].

2.3.1. *Formulació.* El model que utilitzem és una modificació del model logístic. Si l'escala té $K + 1$ nivells, el mètode necessita K models on els paràmetres de les covariables són iguals en tot els models i la única diferència és la constant de cada model. Així, el model es pot expressar com

$$c_k(x) = \ln \left[\frac{P(Y \leq k|x)}{P(Y > k|x)} \right] = \ln \left[\frac{\phi_0(x) + \phi_1(x) + \dots + \phi_k(x)}{\phi_{k+1}(x) + \phi_{k+2}(x) + \dots + \phi_K(x)} \right] = \tau_k - x^T \beta$$

per $k = 0, \dots, K - 1$ on $P(Y = k|x) = \phi_k(x)$.

La interpretació dels paràmetres es fa de manera idèntica que en el model logístic. Es pot comprovar fàcilment que l' OR dels paràmetres són $OR(x_j + c, x_j) = \exp(c\beta_j)$ per a les contínues i $OR(x_{i,l}) = \exp(\beta_{i,l})$ per a les categòriques.

2.3.2. *Estimació.* Per a estimar el model utilitzarem el mètode de la màxima versemblança. Es maximitzarà la versemblança respecte els paràmetres del model. La versemblança del model és

$$l(\beta) = \prod_{i=1}^n [\phi_0(x_i)^{z_{0i}} \cdot \dots \cdot \phi_K(x_i)^{z_{Ki}}].$$

Computacionalment, és millor maximitzar el logaritme de la versemblança,

$$L(\beta) = \sum_{i=1}^n [z_{0i} \ln(\phi_0(x_i)) + \dots + z_{Ki} \ln(\phi_K(x_i))].$$

La maximització es fa derivant respecte dels paràmetres i igualant a zero. Com en el model logístic tenim les equacions de la versemblança. En aquest cas, són $K + p$ equacions. La solució del sistema d'equacions ens dona els paràmetres del model. Per calcular la variància dels paràmetres es fa de manera anàloga al model logístic.

La inferència sobre el model es fa d'igual manera que en el model logístic estàndard. Les mesures de bondat es poden generalitzar per poder utilitzar-les amb els models de regressió logística ordinal. A part, haurem de comprovar l'assumpció de que l' OR és constant en els diferents models amb un test d'igualtat dels paràmetres.

Capítol 3

Resultats

En aquesta capítol, desenvoluparem l'anàlisi principal. Començarem descrivint les variables d'estudi. Després, aplicarem els mètodes a cada problema i explicarem els resultats obtinguts.

1. Descriptiva general

Aquesta primera descriptiva està enfocada en la distribució de les diverses variables en la visita basal. L'anàlisi de cada problema comença amb la descriptiva específica per al problema.

Les variables d'ajust de l'estudi són l'edat, el sexe, la raça i el NIHSS. La taula 1 conté la descripció de les variables d'estudi i d'ajust.

	Variable		Mitjana	Desviació	Mediana	IQR	
	Edat		67,73	12,33	71	17	
	NIHSS		14,16	4,94	13	7	
	Tractament		Sexe		Raça		
	Placebo	Citicolina	Homes	Dones	Blanca	Negra	Altres
N	446	453	470	429	718	134	47
%	49,6	50,4	52,7	47,2	79,9	14,9	5,2

TAULA 1. Taula basal amb la distribució de les variables d'estudi i d'ajust.

El tractament divideix els pacients en dues mostres iguals, ja que l'assaig clínic té un disseny balancejat pel tractament (1 a 1). En la mostra tenim més homes que dones, però les proporcions són pròximes al 50%. La raça de cada pacient no és una variable d'ajust balancejada, perquè en la mostra hi ha un 80% de pacients de raça blanca, un 15% de pacients de raça negra i 5% d'altres races.

La mitjana d'edat dels participants era de 67,73 anys ($s = 12,33$ anys) i una mediana de 71 anys ($IQR = 17$ anys). La diferència entre la mitjana i la mediana ens indica que la distribució no és simètrica (figura 1). Tot i tenir pacients de 18

a 85 anys, la possibilitat de tenir un AVC isquèmic és més alta com més vell és el pacient. En la mostra només un 10% dels pacients té menys de 51 anys i la meitat dels pacients té més de 70 anys. La distribució creixent és típica dels pacients amb un AVC.

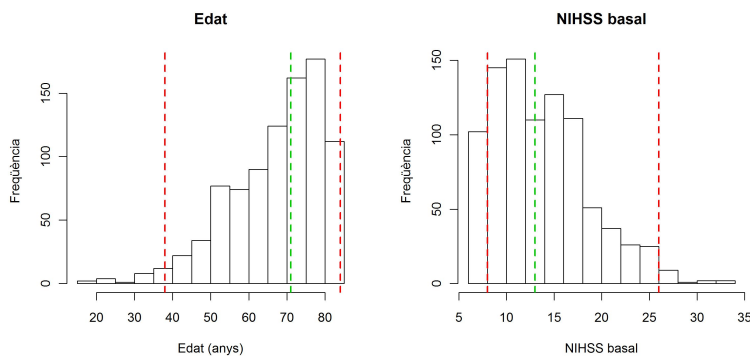


FIGURA 1. L'histograma de l'esquerra correspon a l'edat i el de la dreta, correspon al NIHSS basal. La línia verda representa la mediana de la distribució i les vermelles contenen el 95% de la distribució empírica.

L'índex NIHSS mesura les seqüeles provocades per un AVC. La mitjana del NIHSS basal és de 14,16 ($s = 4,94$) i la seva mediana de 13 ($IQR = 7$). La figura 1 presenta l'histograma de l'índex. L'histograma té tots els valors superiors a 7, ja que hi han alguns pacients que no complien la condició de tenir un NIHSS basal superior o igual a 8. A més, una de les condicions de l'assaig clínic era que no podien haver-hi més de 4 pacients amb un índex superior a 18 en cada centre. Aquesta condició, ha provocat que un 75% dels pacients tenen un índex inferior a 18 i que la distribució del NIHSS s'acumuli en els valors inferiors de l'índex NIHSS.

La descriptiva necessària per cada problema s'explicarà per separat i juntament amb la solució proposada.

2. Hipotèrmia induïda pel tractament

El primer anàlisi està centrat en respondre si el tractament produeix una disminució de la temperatura corporal. Començarem descrivint l'evolució de la temperatura segons el temps de seguiment, del tractament i de les variables d'ajust. Finalment, presentarem l'efecte que té la citicolina sobre l'evolució de la temperatura corporal.

2.1. Descriptiva. Comencem la descriptiva fixant-nos en la distribució conjunta de totes les mesures de la temperatura. La figura 2 presenta l'histograma de la temperatura. Aquesta distribució té 4.089 mesures repartides per pacient i visites. La mitjana global de la temperatura és de $36,62^{\circ}\text{C}$ ($s = 0,56^{\circ}\text{C}$) i la mediana de $36,61^{\circ}\text{C}$ ($IQR = 0,72^{\circ}\text{C}$). La temperatura de persones sanes està dins de l'interval $36,5 \pm 0,20^{\circ}\text{C}$. Així, els pacients de l'estudi tenen una temperatura lleugerament superior; però amb molta més variabilitat que les persones sanes.

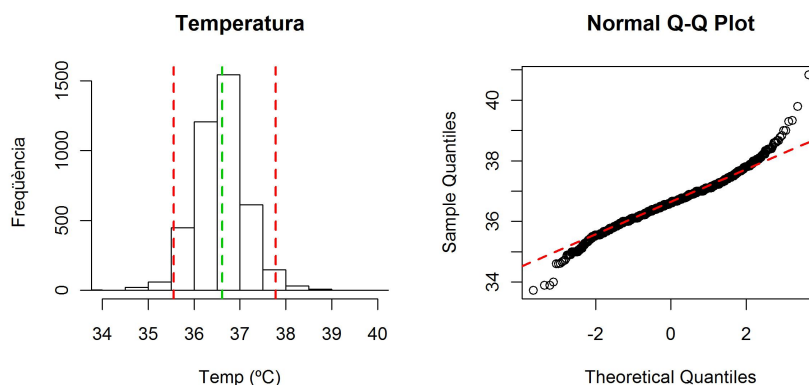


FIGURA 2. La figura de l'esquerra és l'histograma de totes les mesures de la temperatura. La línia verda marca la mediana de la mostra i les línies vermelles contenen el 95% de la mostra. La figura de la dreta és el QQ-Normal de les temperatures.

El rang global de la temperatura corporal està comprès entre $[33, 72; 40, 83]$, però el 95% de la distribució està continguda a $[35, 56; 37, 78]$. Hi han 66 mesures en 63 pacients que es consideren en estat de febre (a partir de 38°C) o hipotèrmia (per sota de 35°C). A més, l'interval del 95% dels casos segons una normal és $[35, 50; 37, 74]$ que és molt semblant a l'interval previ que conté el 95% de la distribució empírica.

Com la temperatura va estar mesurada en cada visita, pot tenir una evolució temporal. A més, l'evolució temporal pot variar segons el tractament o les variables d'ajust. En la següent descripció ens centrarem en l'evolució temporal de la temperatura corporal segons diversos determinants.

La primera descripció és l'evolució temporal de la temperatura corporal (figura 3 i taula 2). Durant les tres primeres visites la temperatura disminueix i és menys

variable. A partir, de la tercera visita la distribució es manté estable. En l'anàlisi inclourem un efecte temporal per mesurar la disminució.

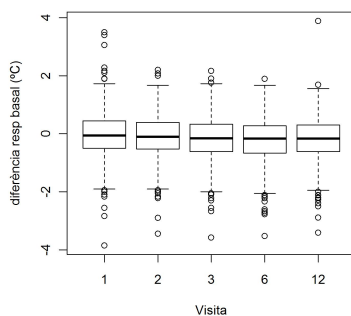


FIGURA 3. Boxplot entre la visita setmanal (temps) i la temperatura corporal.

Visita setmanal	N	missing	Mínim	Mediana	IQR	Màxim	Mitjana	Variància
1	492	407	-3,8556	-0,0556	0,9444	3,5000	-0,0371	0,8325
2	703	196	-3,4444	-0,1000	0,9167	2,2000	-0,0917	0,7308
3	657	242	-3,5778	-0,1556	0,9444	2,1667	-0,1685	0,7726
6	638	261	-3,5222	-0,1667	0,9444	1,8889	-0,2234	0,7582
12	593	303	-3,4111	-0,1667	0,9111	3,8889	-0,1927	0,7523

TAULA 2. Descripció de la temperatura segons la visita.

S'observa una tendència global lleugerament decreixent, però no s'ha d'oblidar la pròpia evolució del pacient, ja que l'evolució pot canviar pacient a pacient. En la figura 4, hi han les trajectòries de 48 pacients. La majoria de trajectòries personals són pròximes a ser lineals amb pendent pròxim a 0, però també hi han alguns pacients amb corbes molt fluctuants i no-lineals o amb pocs valors mesurats.

A continuació, estudiarem si existeix la relació entre la temperatura i les variables d'ajust i de quina manera és.

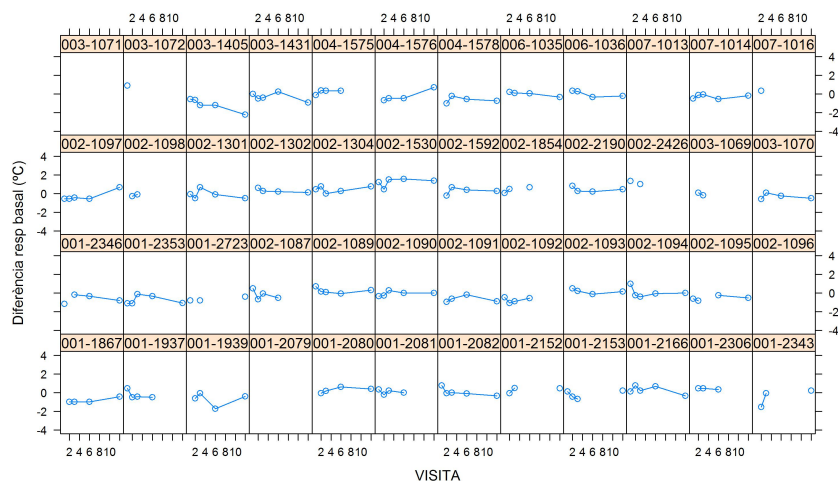


FIGURA 4. Evolució de la temperatura segons el número de visita per a diversos pacients. Les línies uneixen les mesures consecutives.

La primera relació que estudiem és entre l'evolució temporal i el tractament. A la taula 3 i a la figura 5, hi ha una descripció numèrica i gràfica d'aquesta relació. Els dos tractaments tenen un efecte similar, ja que les distribucions per visita són indistingibles, visualment parlant. Les poques diferències són que el grup de tractats té una mitjana de temperatura basal lleugerament inferior al grup control i que en les dues primeres visites el grup control té una mica més de variabilitat.

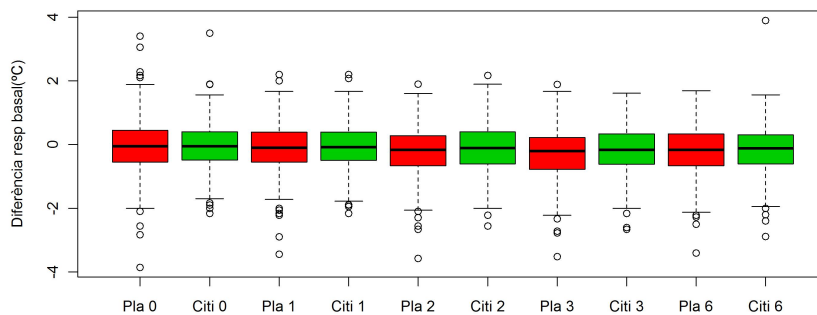


FIGURA 5. Boxplot de l'evolució temporal segons el tractament. El color vermell representa les visites del grup tractat i el verd al grup control.

Visita	Tractament	N	Missing	Mínim	Mediana	IQR	Màxim	Mitjana	Variància
1	Placebo	238	208	-3,8556	-0,0556	1,0000	3,4000	-0,0358	0,8929
	Citicolina	254	199	-2,1667	-0,0556	0,8750	3,5000	-0,0383	0,7735
2	Placebo	343	103	-3,4444	-0,1000	0,9444	2,2000	-0,1187	0,7571
	Citicolina	360	93	-2,1667	-0,0778	0,8889	2,2000	-0,0660	0,7050
3	Placebo	317	129	-3,5778	-0,1667	0,9444	1,9000	-0,2055	0,7912
	Citicolina	340	113	-2,5556	-0,1111	1,0028	2,1667	-0,1341	0,7543
6	Placebo	305	141	-3,5222	-0,2000	1,0000	1,8889	-0,2753	0,7955
	Citicolina	333	120	-2,6667	-0,1667	0,9556	1,6111	-0,1759	0,7203
12	Placebo	286	153	-3,4111	-0,1667	1,0000	1,6889	-0,2290	0,7705
	Citicolina	300	153	-2,8889	-0,1167	0,9111	3,8889	-0,1573	0,7337

TAULA 3. Descripció de l'evolució de la temperatura segons el tractament.

La primera variable d'ajust que estudiem és el sexe del pacient. La figura 6 i la taula 4 presenten una descripció visual i numèrica de l'evolució de la temperatura segons el sexe del pacient. L'evolució en els dos grups és semblant. Només destaca que en les dues últimes visites els homes tenen una temperatura més baixa.

Visita	Sexe	N	Missing	Mínim	Mediana	IQR	Màxim	Mitjana	Variància
1	Dona	240	189	-3.8556	-0.1111	1.0000	3.4000	-0.0817	0.8314
	Home	249	218	-2.8333	0.0000	0.9917	3.5000	0.0054	0.8330
2	Dona	337	109	-3.4444	0.0000	0.8889	2.2000	-0.0678	0.7577
	Home	366	104	-2.1667	-0.1111	0.9333	2.2000	-0.1137	0.7055
3	Dona	319	127	-3.5778	-0.1667	0.9722	2.1667	-0.1995	0.7741
	Home	338	132	-2.5556	-0.1111	0.9917	1.9000	-0.1394	0.7712
6	Dona	309	137	-3.5222	-0.1333	0.9444	1.6111	-0.1867	0.7667
	Home	329	141	-2.7222	-0.2222	0.9556	1.8889	-0.2579	0.7497
12	Dona	283	163	-3.4111	-0.1111	0.9167	3.8889	-0.1431	0.7859
	Home	310	160	-2.8889	-0.1833	0.9528	1.3333	-0.2380	0.7186

TAULA 4. Descripció de l'evolució de la temperatura per sexe.

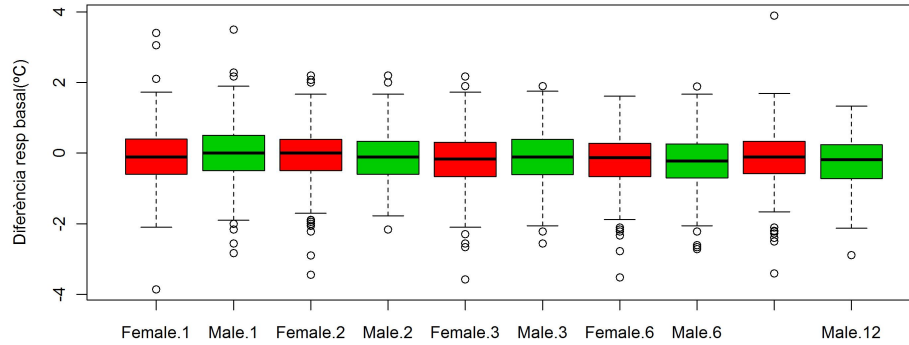


FIGURA 6. Boxplot de l'evolució temporal segons el sexe. El color vermell representen les visites de les dones i el verd als homes.

La segona variable d'ajust més important és la raça del pacient, encara que esperem que l'evolució no depengui de la raça. El problema d'aquesta variable d'ajust és que els pacients de raça blanca representa el 80% dels pacients. Es pot esperar que si la descriptiva mostra un efecte, aquest no serà significatiu en el model que construïm. En la taula 5 i la figura 7 es presenta la descriptiva de l'evolució segons la raça. La descriptiva no destaca cap efecte.

Visita	Raça	N	Missing	Mínim	Mediana	IQR	Màxim	Mitjana	Variància
1	Blanca	376	342	-3.8556	-0.0556	0.9111	3.5000	-0.0623	0.8163
	Negra	87	47	-2.1000	0.0000	1.0556	3.0556	0.0497	0.9416
	Altres	30	17	-1.1111	-0.0556	0.7917	1.6111	0.0270	0.6911
2	blanca	559	159	-3.4444	-0.1000	0.9167	2.2000	-0.1040	0.7320
	Negra	109	25	-1.9444	0.0556	0.7778	1.6667	-0.0410	0.7102
	Altres	35	12	-2.0000	-0.1111	1.0056	2.2000	-0.0533	0.7864
3	Blanca	517	199	-3.5778	-0.1111	0.9444	2.1667	-0.1638	0.7837
	Negra	102	32	-1.8889	-0.1944	0.9472	1.9000	-0.1625	0.7641
	Altres	36	11	-1.8889	-0.2222	0.6500	1.0556	-0.2534	0.6359
6	Blanca	506	212	-3.5222	-0.1833	0.9694	1.8889	-0.2387	0.7616
	Negra	102	32	-2.1667	-0.1111	1.0528	1.5000	-0.1368	0.7371
	Altres	30	17	-2.1111	-0.2222	0.6583	1.4444	-0.2611	0.7765
12	Blanca	468	250	-3.4111	-0.1667	1.0000	1.5556	-0.2193	0.7584
	Negra	94	76	-2.1111	0.0000	0.8667	3.8889	-0.0463	0.7752
	Altres	31	16	-1.3333	-0.1667	0.6111	0.7778	-0.2351	0.5215

TAULA 5. Descripció de l'evolució de la temperatura per raça.

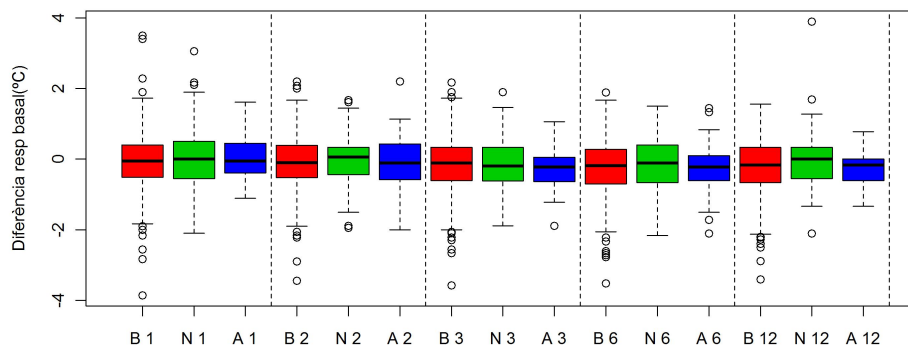


FIGURA 7. Boxplot de l'evolució temporal segons la raça. El color vermell representen les visites de la raça blanca, la verda representen els asiàtics, la blava representa els negres, el cian representa els hispans i la lila les altres races. Les línies verticals separen les diferents visites.

Les últimes descripcions que farem estan centrades en la influència de les variables numèriques sobre l'evolució de la temperatura. La descriptiva la farem amb un bivariat estratificat per visites.

En la figura 8, hi ha representada la relació entre la temperatura i l'edat per cada visita. Visualment, no observem cap tendència global dintre de cada visita i tampoc sembla que entre visites l'edat tingui efecte.

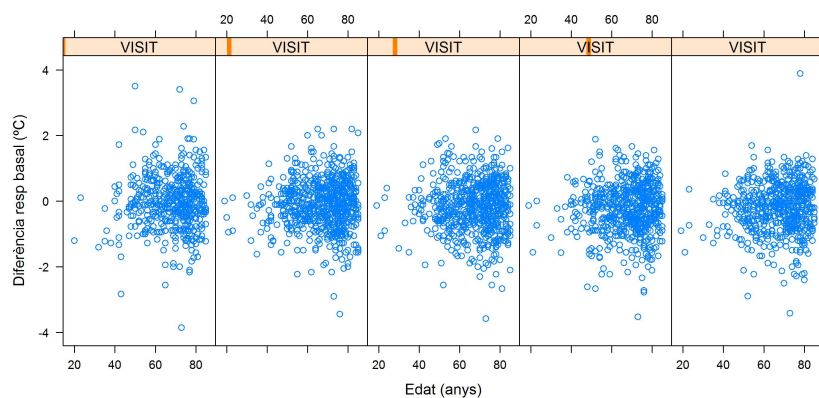


FIGURA 8. Diagrames bivariats entre l'evolució de la temperatura corporal i l'edat del pacient segons la visita.

L'altre variable numèrica és el NIHSS basal. La figura 9 presenta els bivariats per cada visita entre la temperatura i el NIHSS. En aquest cas, tampoc s'observa un efecte del NIHSS sobre la temperatura ni sobre l'evolució. En el model, s'espera que no tingui cap efecte.

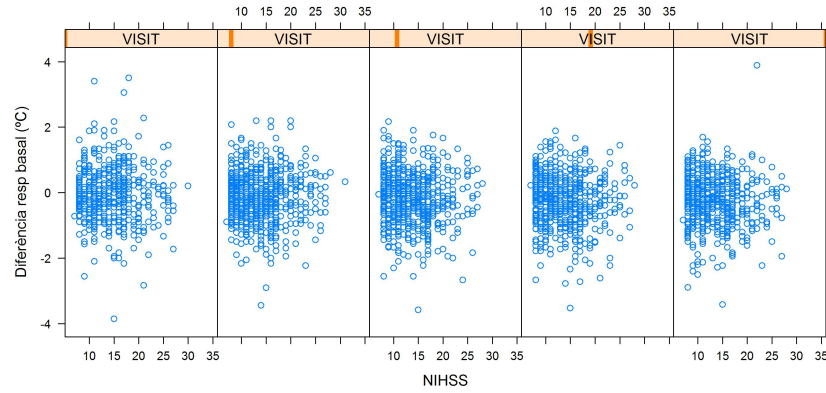


FIGURA 9. Diagrames bivariats entre l'evolució de la temperatura corporal i el NIHSS basal segons la visita.

2.2. Models. Per a quantificar l'efecte de la citicolina sobre l'evolució temporal construïrem dos models. Un primer sense la citicolina per reproduir l'evolució de la temperatura amb les variables d'ajust. En el segon introduïrem l'efecte de la citicolina en el model previ i quantificarem l'efecte del fàrmac.

En els models, l'evolució de la temperatura corporal no vindrà definida per la visita setmanal. Utilitzarem el temps de seguiment en dies. El temps de seguiment mesura quants dies ha passat des de l'admissió a urgències del pacient. El temps de seguiment és una mesura temporal més precisa que la visita setmanal, ja que la visita setmanal no es va fer en el mateix dia de la setmana en tots els pacients. La figura 10 presenta l'evolució de la temperatura segons el dia de seguiment de 48 pacients.

2.2.1. *Model sense el tractament.* A l'annex, hi ha la construcció del model que ajusta millor l'evolució de la temperatura amb les variables d'ajust: temps de seguiment, sexe, raça, NIHSS basal i edat. En aquest apartat, resumirem la construcció del model.

El model per a l'evolució de la diferència de temperatura respecte la basal ($dT_p(t)$) estarà format per quatre components que varien segons el pacient (p). La primera modela l'evolució temporal segons el dia de seguiment ($f_p(t)$). La segona modifica la temperatura segons les variables d'ajust ($g_p(t)$). La tercera inclou la interacció entre el temps de seguiment i les variables d'ajust ($i_p(t)$). I la quarta modela els residus segons un model lineal ($\epsilon_p(t)$). El model es pot descriure com

$$dT_p(t) = f_p(t) + g_p(x) + i_p(t, x) + \epsilon_p(t)$$

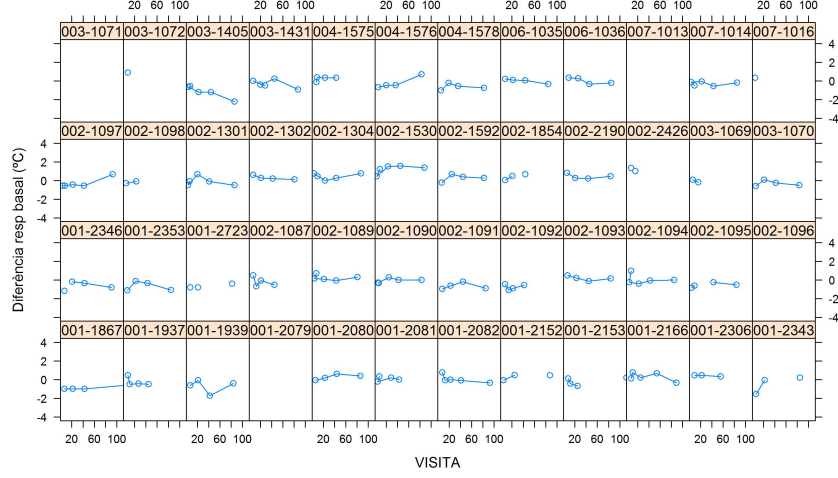


FIGURA 10. Evolució de la temperatura segons el dia de seguiment per a diversos pacient. Les línies uneixen les mesures consecutives.

on t representa el dia de seguiment, x són les variables d'ajust i p representa el pacient.

Com que només ens interessa descriure l'evolució temporal dins de l'interval d'estudi, pensem que l'ajust per un polinomi és suficientment bona. La millor aproximació a l'evolució temporal és un polinomi quadrat. Així,

$$f_p(t) = C_{p,0} + C_{p,1} \cdot t + C_{p,2} \cdot t^2$$

on $C_{p,0}$, $C_{p,1}$ i $C_{p,2}$ poden ser diferents per a cada pacient.

Només hem trobat que l'evolució es modifica segons el sexe del pacient. Tant homes com dones tenen una temperatura basal iguals, però durant el seguiment les seves components lineals són diferents. Per tant,

$$i_p(t, x) = S \cdot [Sexe_p = Home] \cdot t$$

on el paràmetre S té un efecte fix per a tots els pacients.

Per últim, només queda la component dels residus. De l'observat a la descriptiva, sabem que la variància es reduïa a l'augmentar la visita. Introdueixim una funció de variàncies que ens permeti reduir la variància segons el temps de seguiment, com la funció variància potència del temps de seguiment. Així,

$$Var[\epsilon_p(t)] = \sigma^2(\delta_1 + |t|^{\delta_2})^2$$

on σ^2 és la variància intra-individu, δ_1 és l'asíptota horitzontal de la variància (variància màxima o mínima) i δ_2 mesura la velocitat del creixement ($\delta_2 > 0$) o del decreixement ($\delta_2 < 0$) de la variància respecte el dia de seguiment.

Per tant, el model final és

$$dT_p(t) = [C_{p,0} + C_{p,1} \cdot t + C_{p,2} \cdot t^2] + [S \cdot [Sexe_p = Home] \cdot t] + \epsilon_p(t)$$

$$\text{on } \text{Var}[\epsilon_p(t)] = \sigma^2(\delta_1 + |t|^{\delta_2})^2.$$

El model estimat està a la taula 7. Globalment, tenim una tendència quadràtica concava, però la curvatura de l'aproximació és poc pronunciada. Això, és degut al llarg rang del seguiment dels pacients, $[0, 100]$ dies. Tots els pacients no tenen una tendència quadràtica concava, ja que l'efecte aleatori sobre el terme quadràtic pot canviar el signe d'aquest paràmetre.

Paràmetre	Estimació	Efectes fixos			Efectes aleatoris			Funció de Variància
		Error Estimació	p -valor de Wald	p -valor ANOVA	Desviació Estàndard	Correlació C_0	C_1	
C_0	-0,011474	0,030089	0,7030	>, 0001	0,624463			0,282357 1,019791 -0,197811
C_1	-0,005468	0,001354	0,0001	>, 0001	0,015494	-0,240		
C_2	0,000050	0,000013	0,0001	0,0001	0,000125	0,184	-0,979	
S	-0,001633	0,000598	0,0063	0,0063				
σ								
δ_1								
δ_2								

TAULA 6. Estimació del model sense el tractament.

La variabilitat de la constant del model remarca que l'extrapolació de la temperatura basal segons l'evolució de les temperatures sol no coincidir amb la mesurada. La variància intra-individu és la segona font de variància més important. A més, la variància intra-individu al principi de l'estudi és superior a la del final de l'estudi ($\text{Var}(\epsilon_p(1)) = 0.325245 > \text{Var}(\epsilon_p(91)) = 0.162920$). El decreixement de la variància és coherent amb l'observat en la descriptiva.

La figura 11 conté les aproximacions de 18 pacients. L'aproximació és adequada pels pacients amb totes les mesures i amb una evolució suau. En canvi, pels pacients amb poques mesures o evolucions abruptes el model no ajusta de manera correcta. Per a poder predir els canvis abruptes necessitariem una variable externa explicant si va ocórrer algun fenomen i quan, però en la base de dades no es va recollir.

2.2.2. Validació del model sense tractament. Per considerar que el model és correcte necessitem validar-lo. La validació es farà sobre els supòsits que hem fet sobre la distribució dels residus del model. Estudiarem les assumpcions del model amb els gràfics de validació del model lineal mixt.

La primera suposició a comprovar és que el model és capaç d'ajustar-se a les dades d'estudi. Per a comprovar-ho, estudiarem el gràfic bivariat entre els valors estimats (eix OX) i els residus estandarditzats del model (eix OY), figura 12. El model és capaç d'ajustar-se als valors centrals de les mesures observades i només el 1,67% dels residus estandarditzats estan fora de l'interval $[-1,96; 1,96]$; però els valors baixos ($dT < -2$) estan sobreestimats i els alts ($dT > 1$) infraestimats.

La segona suposició és que els residus intra-individu es distribueixen com una normal. La figura 13 presenta el QQ-plot dels residus del model segons una normal estàndard. La distribució del residus del model, llevat d'alguns atípic en els extrems,

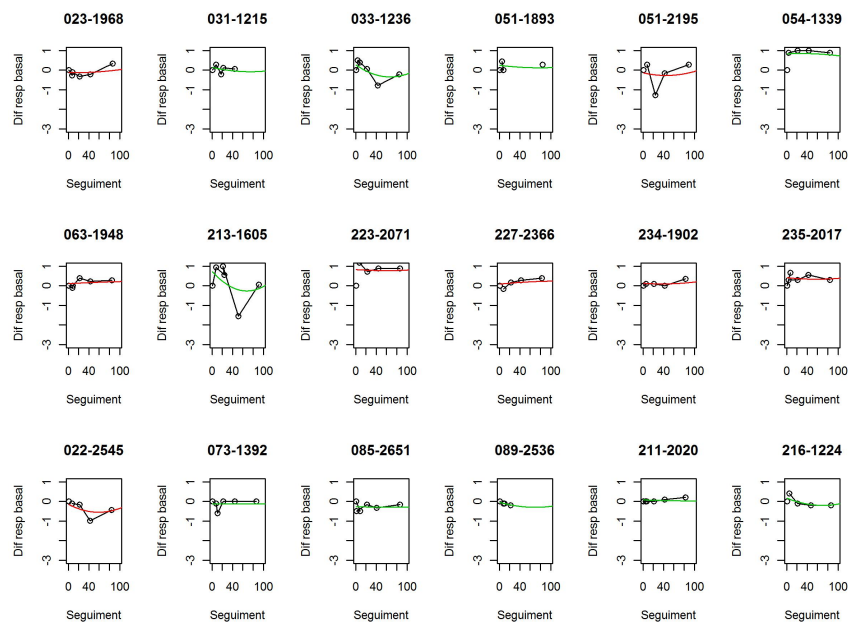


FIGURA 11. Aproximació quadràtica, sense el tractament, de l'evolució de la temperatura per a 18 pacients. Les línies i punts negres corresponen a les dades del pacient. Les línies vermella i verda corresponen a l'aproximació quadràtica de les dones i dels homes, respectivament.

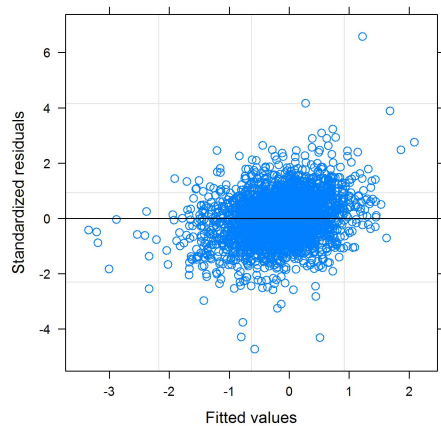


FIGURA 12. Bivariat entre els valors reals i els residus estandarditzats.

es distribueix com una normal. Ja que la part central del diagrama bivariat és pròxim a la recta bisectriu. Així, l'assumpció de normalitat dels residus intra-individu es una suposició coherent.

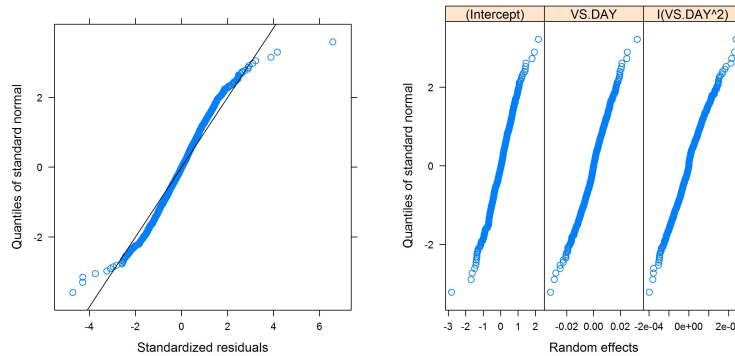


FIGURA 13. QQ-normals dels residus del model i dels tres efectes aleatoris.

La última suposició és que els efectes aleatoris es distribueixen com una normal. La figura 13 presenta els tres QQ-plots dels efectes aleatoris enfront d'una normal estàndard. L'efecte aleatori de la constant és pròxima a una distribució normal. Les distribucions dels efectes lineals i quadràtics són poc variables i el seu ajust a una normal no està clar.

El model no és capaç de reproduir totalment les dades reals, queda una font de variació deguda a esdeveniments no modelables. Amb les variables d'ajust disponibles no en sobra o falta cap d'elles. Per tant, no es pot millorar amb les variables disponibles.

2.2.3. *Model amb tractament.* Un cop hem construït el model sense el tractament, contrastarem si l'efecte de la citicolina és significatiu. Esperem que la citicolina provoqui una disminució de la temperatura corporal més intensa que els pacients que han rebut el placebo. Per modelar la temperatura corporal segons el tractament farem tres suposicions. La primera és que el tractament no influeix sobre el nivell de l'evolució. La segona és que el tractament modifica el comportament de la relació entre la temperatura i el dia de seguiment. La tercera és que el tractament influeix de la mateixa manera a tots els pacients. Per tant, la citicolina entrarà en el model dins la component $i_p(t)$ i es descriu com $i_p(t) = LT \cdot [TREAT_p = Citicolina] \cdot t + QT \cdot [TREAT_p = Citicolina] \cdot t^2$.

L'efecte de la citicolina sobre l'evolució de la temperatura corporal és no significatiu, si suposem que el model construït sense el tractament és un model complet i correcte. A més, donada la potència de l'estudi, podem concloure que la Citicolina no té efecte en l'evolució de la temperatura.

Paràmetre	Efectes fixos				Efectes aleatoris			Funció de Variància
	Estimació	Desviació Estàndard	p -valor de Wald	p -valor ANOVA	Desviació Estàndard	Correlació C_0	C_1	
C_0	-0,011071	0,030098	0,7130	>, 0001	0,624783			
C_1	-0,006626	0,001760	0,0002	>, 0001	0,015542	-0,243		
C_2	0,000061	0,000017	0,0004	0,0001	0,000125	0,188	-0,980	
S	-0,001620	0,000598	0,0068	0,0064				
LT	0,002181	0,002122	0,3043	0,7709				
QT	-0,000022	0,000022	0,3245	0,3245				
σ								0,280953
δ_1								1,025890
δ_2								-0,196495

TAULA 7. Estimació del model amb el tractament.

3. La mort en funció de la temperatura basal

En el segon anàlisi, estudiarem la relació entre la mort i la temperatura basal. Començarem fent la descriptiva de la mort segons la temperatura basal, el tractament i les variables d'ajust. Per contrastar estadísticament els resultats previs, construirem un model de regressió logística per quantificar l'efecte no ajustat i ajustat.

3.1. Descriptiva.

Per la descriptiva, construirem una taula de contingència per descriure la relació entre la mort i el tractament o les variables d'ajust categòriques. Comprovarem la relació amb el test de χ^2 per a les taules de contingència. Per a les variables contínues descriurem la distribució segons si estan vius o morts. A més, amb el t -test comprovarem si les mitjanes de cada grup són iguals.

La taula de contingència està representada en la taula 8. En tota la mostra s'ha mort un 17% dels pacients. Si ho mirem per a cada grup, tenim la mateixa mortalitat en cada grup, un 17%. Per a les variables sexe i raça, les proporcions de vius i morts per cada categoria són pròximes i, a més, el test χ^2 no rebutja les hipòtesis nul·les.

		Viu n (%)	Mort n (%)	Total n	OR	Test de χ^2 p -valor
Tractament	Citicolina	376 (83,0%)	77 (17,0%)	453	Ref	1
	Placebo	370 (83,0%)	76 (17,0%)	446	0,99	
Sexe	Dones	349 (81,3%)	80 (18,7%)	429	Ref	0,2481
	Homes	397 (84,5%)	73 (15,5%)	470	0,80	
Raça	Blanca	597 (83,1%)	121 (16,9%)	718	Ref	0,9155
	Negra	108 (80,6%)	26 (19,4%)	134	1,54	
	Altres	41 (87,2%)	6 (12,8%)	47	0,72	
Total		746 (83,0%)	153 (17,0%)	899		

TAULA 8. Taula de contingència per a les variables categòriques amb el test χ^2 .

La taula 9 presenta les distribucions dels vius i els morts per a cada variable numèrica. En la figura 14, hi han els corresponents boxplots. La distribució de la temperatura basal per les dues categories són visualment indistingibles. En canvi, les distribucions del NIHSS basal i de l'edat són visualment separables. El grup dels morts mostra una distribució amb valors més grans que el grup dels vius. En el model contrastarem si els valors elevats del NIHSS basal o de l'edat porten a un augment de la mortalitat. Aquests resultats queden validats pel t -test (taula 10) per a cada variable segons la mort del pacient, ja que només no es rebutja la hipòtesis nul·la per a la temperatura basal.

	Estat	N	mis	Mínim	Mediana	IQR	Màxim	Mitjana	Variància
Edat	Viu	746	0	19	70	18	85	66,68	159,6083
	Mort	153	0	37	75	12	85	72,23	90,8486
Temperatura basal	Viu	708	38	35,11	36,70	0,8111	39,80	36,73	0,3566
	Mort	142	11	34,60	36,78	0,9556	39,00	36,73	0,5406
NIHSS basal	Viu	746	0	7	12,5	6	32	13,43	20,8919
	Mort	153	0	9	17,0	7	34	17,76	26,2767

TAULA 9. Descripció de la variable mort segons l'edat, la temperatura basal i el NIHSS basal.

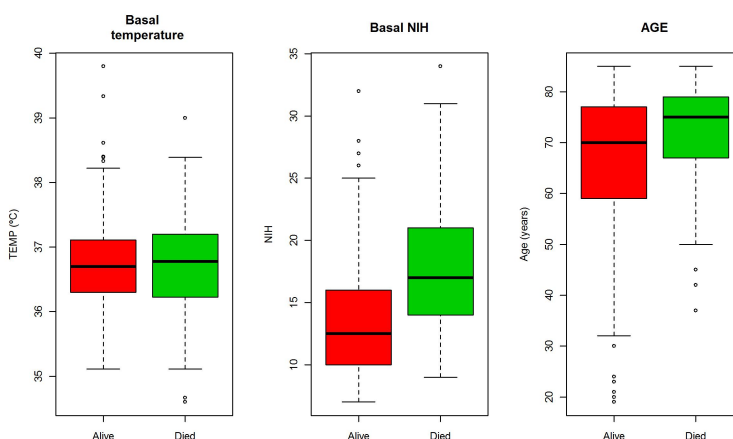


FIGURA 14. Boxplot per a cada variable numèrica segons si acaben l'estudi vius o morts. El color vermell representen els pacients vius i el color verd els morts.

Variable	Mitjana en els vius	Mitjana en els morts	<i>t</i> -test <i>p</i> -valor
Temperatura basal	36,7280	36,7279	0,9992
NIHSS basal	13,4263	17,7582	< 0,0001
Edat	66,8070	72,2288	0,0001

TAULA 10. *t*-test per les variables numèriques segons la mort del pacient.

3.2. Models. Per quantificar el risc de major mortalitat calcularem l'odds ratio (OR). Per calcular els OR utilitzarem el model de regressió logística. El model per als OR no ajustat contindrà només la variable d'estudi i per als ajustats totes les variables d'estudi. Els OR ajustats mesuren l'efecte de la variable d'estudi tenint en compte l'efecte de les altres variables.

Abans de construir el model, és necessari comprovar la relació de linealitat entre el logit i les variables quantitatives. Per a la comprovació, dividirem la variable

quantitativa en 10 categòriques indicadores del decil que es troba el valor. Amb les variables calcularem els OR de la mort pel canvi del decil i al $i + 1$. Compararem els diferents OR amb un forest plot, figura 15. Les taules que contenen els diferents OR per a la temperatura basal, el $NIHSS$ basal i l'edat estan als annexs.

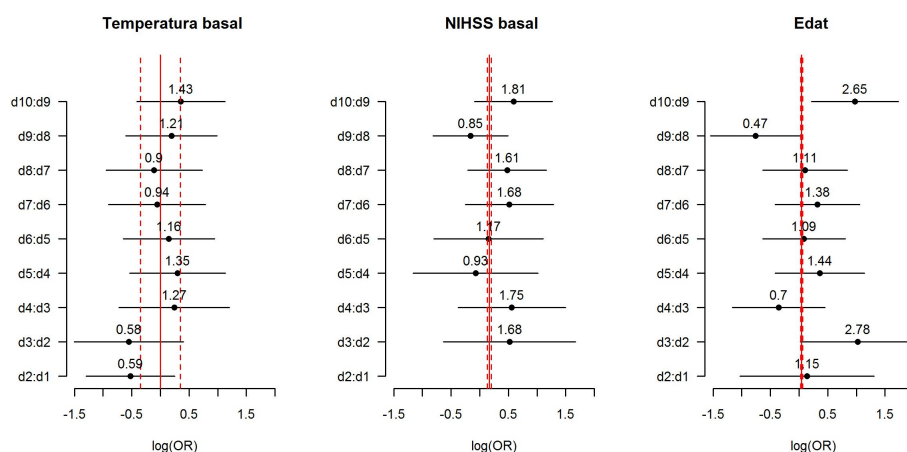


FIGURA 15. Foresplot del $\ln(OR)$ del canvi de decil per a cada variable quantitativa. Els valors són els OR pel canvi de decil i a $i + 1$. La línia vermella contínua marca el valor del $\ln(OR)$ sota el supòsit de linealitat i les discontinües el seu interval de confiança al 95%.

La temperatura i el NIHSS basal tenen, visualment, uns OR equivalents al canviar de decil. Per tant, simplifiquem el problema i suposarem que tenen un efecte lineal, ja que no observem cap patró que suggereixi un model alternatiu. Amb l'edat, no concloem, visualment, que els diferents OR siguin equivalents, ja que tenim 7 OR equivalents i dos de diferents (el 2n i 8è començant per a baix). Però traient aquests dos OR sí que té un comportament equivalent i superior al estimat conjuntament. Així, simplifiquem i suposarem que l'edat té un efecte lineal sobre els logits i l'inclourem de manera quantitativa.

Els OR no ajustats i ajustats estan a la taula 11. L'augment d'una unitat la temperatura basal té associat un OR no ajustat de 0,9998 i ajustat de 0,9149. Però el seu efecte és no significatiu, perquè els seus intervals de confiança inclouen el 1. Per tant, si el model és correcte, l'augment de la temperatura basal té un efecte no significatiu.

De les variables d'ajust del model, les úniques significatives són les numèriques, el NIHSS basal i l'edat del pacient. Els OR no ajustat i ajustat del NIHSS basal destaquen que l'augment d'una unitat suposa un augment del 18,78% o 19,36% de l'odds de morir, respectivament. L'augment d'una any d'edat comporta l'augment d'un 4,44% o 4,06% de l'odds de morir depenen de si l' OR està no ajustat o ajustat.

Per tant, si el model és correcte, concloem que l'augment d'un grau en la temperatura basal no suposa un augment de l'odds de morir. Tampoc el tractament en

Variable	Categoria	OR	CI (OR,95%)	OR	CI (OR, 95%)
		no ajustat		ajustat	
Temperatura basal	1 unitat	0,9998	[0,7486; 1,3353]	0,9149	[0,6749; 1,2401]
Tractament	Placebo	Ref		Ref	
	Citicolina	0,9970	[0,7040; 1,4118]	0,9228	[0,6227; 1,3675]
Sexe	Dona	Ref		Ref	
	Home	0,5662	[0,8022; 1,1364]	1,0814	[0,7239; 1,6154]
Races	Blanca	Ref		Ref	
	Negra	1,1878	[0,7419; 1,9017]	1,5458	[0,9131; 2,6168]
	Altres	0,7220	[0,2999; 1,7385]	0,4183	[0,1448; 1,2086]
NIHSS basal	1 unitat	1,1812	[1,1393; 1,2246]	1,1936	[1,1468; 1,2422]
Edat	1 unitat	1,0444	[1,0263; 1,0629]	1,0406	[1,0199; 1,0616]

TAULA 11. L'odds ratio de morir no ajustat i ajustat pel canvi de categoria respecte la referència per a les variables categòriques i per a l'augment d'una unitat en les numèriques.

aquest cas té un efecte protector sobre la mort per un AVC isquèmic. Els únics efectes significatius sobre la mort són l'augment d'una unitat del NIHSS basal i l'edat. Així, els pacients amb un AVC isquèmic més greu en la visita basal o de més edat tenen més odds de morir. Aquests serien els pacients amb més risc de morir.

Validació del model Els valors dels estadístics per a la validació del model estan a la taula 12. A més, s'ha construït la corba ROC (figura 16) per a visualitzar el comportament de la sensibilitat, la especificitat i l'habilitat del model per a discriminar.

Estadístic de validació	Valor	<i>p</i> -valor
Raó de versemblança	116,970	$\leq 0,0001$
D de Sommers	0,563	
Gamma	0,565	
Tau de Kendall	0,157	
R^2	0,216	
Test de Hosmer-L		<i>p</i> -valor
y=1		0,0033

TAULA 12. Estadístics de validació i capacitat predictiva del model logístic per a la mort del pacient.

La raó de versemblança (taula 12) respecte el model buit té un valor elevat que li correspon un *p*-valor molt petit respecte una χ^2 amb 7 graus de llibertat. Per tant, el test de la raó de versemblança rebutja la hipòtesis que el model complet sigui equivalent al model buit. El test de Hosmer-Lemeshow (taula 12) té un *p*-valor molt baix, indicació de que el model no és correcte. Quan ens fixem en la àrea sota

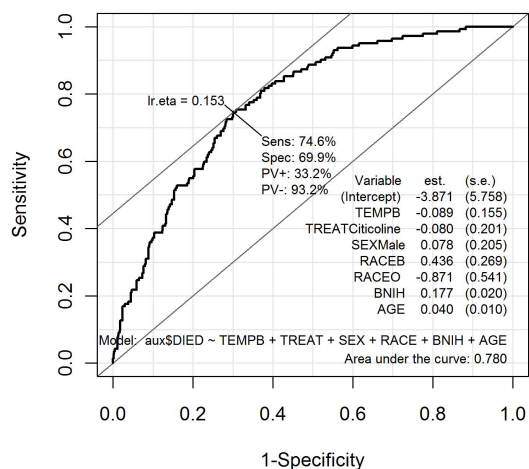


FIGURA 16. Corba ROC per al model logístic de la mort. L'àrea sota la corba ROC és de 0,782.

la corba ROC (figura 16) el model discrimina de manera correcta, ja que tenim una àrea sota la corba de 0,782.

Per tant, es pot afirmar que el model no és perfecte; però discrimina de manera coherent. A més, amb les variables d'estudi no es pot millora el model, ja que una part de la variància de les dades no es pot modelar.

4. La temperatura basal com a factor pronòstic

L'últim anàlisi estudiarem la relació entre l'escala “modified Rankin Scale” (*mRS*) al final de l'estudi i la temperatura basal. L'escala *mRS* mesura la severitat de l'AVC i té valors enters entre 0 i 6. Primer, descriurem el comportament de l'escala *mRS* segons la temperatura basal, el tractament i les variables d'ajust. Construïrem un model de regressió logística ordinal amb odds proporcionals per a tenir l'efecte ajustat de cada variable.

La taula 13 presenta el significat de cada nivell. No utilitzarem l'escala *mRS*, si no una nova categorització. Si els pacients són independents per a activitats bàsiques de la vida diària, serà la categoria $mRS = 0 - 2$; semi-dependents per a activitats bàsiques de la vida diària, $mRS = 3 - 4$; dependents per a les activitats bàsiques de la vida diària, $mRS = 5$; o morts, $mRS = 6$.

Nivell	Significat
0	Asimptomàtic.
1	Sense incapacitat important.
2	Incapacitat lleu de fer alguna de les seves activitats prèvies, que es realitza amb dificultat però sense necessitat ajuda.
3	Incapacitat moderada. Necessita alguna ajuda.
4	Incapacitat moderadament greu. Sense necessitat d'atenció contínua.
5	Incapacitat greu. Totalment dependent, necessitan assistència constant.
6	Mort.

TAULA 13. Significat dels nivells de l'escala *mRS*.

En la descriptiva, construïrem una taula de contingència per a les variables categòriques i calcularem el test χ^2 per a contrastar si existeix una relació entre les variables. Per a les numèriques, descriurem la relació amb boxplots. El model de regressió logística ordinal ens permetrà calcular els OR no ajustats i ajustats.

4.1. Descriptiva. La taula 14 conté la taula de contingència entre les variables categòriques i l'escala *mRS*. El tractament està relacionat amb la resposta de l'escala *mRS*, ja que el test χ^2 rebutja la hipòtesis de no relació amb un p -valor de 0,0292. El test χ^2 rebutja la hipòtesis de no relació amb el sexe, però el p -valor és pròxim a la no significació, p -valor = 0,0486. La relació entre la raça i l'escala és no significativa, és coherent amb els resultats previs on la raça no influa sobre les variables d'estudi.

La figura 17 presenta els boxplots de les variables numèriques segons el valor de l'escala *mRS*. Si ens fixem en la temperatura basal, observem que és distribuïda de la mateixa manera en cada valor de l'escala. Dels diversos boxplots per a la relació entre el NIHSS basal i l'escala, s'observa com la mediana del NIHSS creix de manera lineal amb l'escala, però com més gran és el valor de l'escala *mRS* més variància té el NIHSS basal. En canvi, la relació entre l'edat i l'escala *mRS* no és tan clara, ja que la distribució per cada valor de l'escala és molt variable.

Factor	Escala <i>mRS</i> [n (%)]				Test χ^2 <i>p</i> -valor
	0-2	3-4	5	6	
Tractament					
Citicoline	179 (41,7)	146 (34,0)	30 (7,0)	74 (17,2)	0,0292
Placebo	154 (36,4)	180 (42,6)	17 (4,0)	74 (17,0)	
Sexe					
Dona	152 (37,0)	156 (38,0)	26 (6,3)	77 (18,7)	0.0486
Home	181 (41,0)	170 (38,5)	21 (4,8)	69 (15,6)	
Raça					
Blanca	269 (39,7)	262 (38,6)	31 (4,6)	116 (17,1)	0,2327
Negra	47 (37,0)	46 (36,2)	10 (7,9)	24 (18,9)	
Altres	17 (36,2)	18 (38,3)	6 (12,8)	6 (12,8)	
Total	333	326	47	146	

TAULA 14. Taula de contingència per a cada variable categòrica amb el test χ^2 .

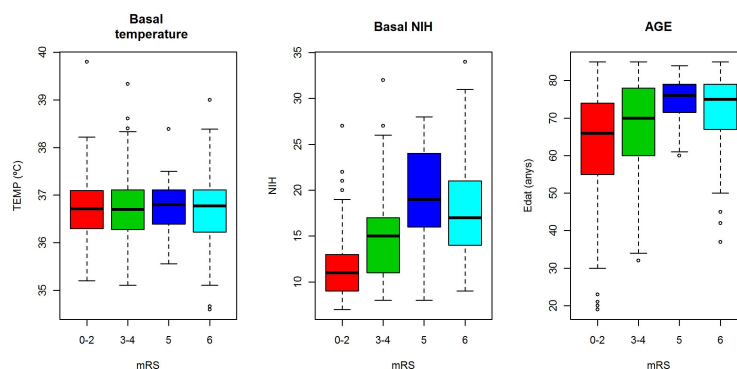


FIGURA 17. Boxplot de cada variable numèrica segons el valor de l'escala *mRS*.

4.2. Model. Per mesurar el valor pronòstic de la temperatura basal sobre les seqüeles, utilitzarem el cumulative odds ratio. El model de regressió logística ordinal ens permet calcular els odds ratio per a la variable resposta ordinal. Estimarem l'efecte tant ajustat com no, per a poder determinar si l'efecte de la variable sola es manté quan introduïm les demés variables d'estudi.

Abans de construir el model, és necessari comprovar la hipòtesis de que els diferents *OR*, es poden aglutinar en un cumulative *OR*. Després s'ha de comprovar la relació de linealitat entre el cumulative logit i les variables quantitatives. Per a la suposició del cumulative *OR*, dividirem la variable quantitativa en 8 categòriques indicadors de l'octil que es troba el valor *i* i calcularem els diversos *OR* de tenir seqüeles pel canvi de l'octil *i* al *i* + 1. Per cada canvi d'octil, construirem 3 models: $mRS \leq 2$, $mRS \leq 4$, $mRS \leq 5$. Compararem els diferents *OR* amb un forest plot, figura 18. Les taules que contenen els diferents *OR* per a la temperatura basal, el NIHSS basal i l'edat, estan a l'annex.

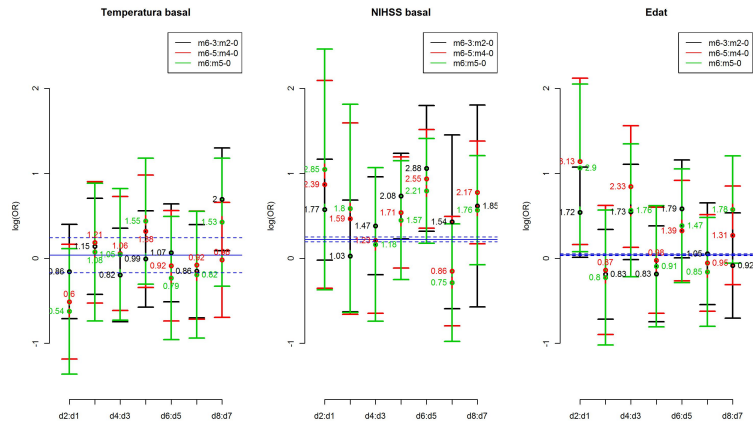


FIGURA 18. Foresplot del $\ln(OR)$ del canvi d'octil i de model per a cada variable quantitativa. Les línies blaves marca el valor del $\ln(OR)$ sota el supòsit de linealitat i el seu interval de confiança al 95%. Els valors mostrats són els diversos OR

La suposició del cumulative OR es satisfà quan els 3 OR per a cada canvi d'octil, són similars. Per tant, com per a les tres variables quantitatives, tenim que els intervals verticals se sobreposen tots i la sobreposició és extensa en cada cas. Pensem que la suposició dels cumulative OR és una bona simplificació.

Per a la suposició de linealitat entre els cumulative OR del canvi d'octil, calcularem els cumulative OR pel canvis d'octil. Si l'assumpció de linealitat és correcta, els diversos cumulative OR tindran valor pròxims. La figura 19 presenta el gràfic dels diversos cumulative OR per al canvi d'octil en cada variable.

Per a la temperatura basal, els OR tenen un comportament constant. Així, suposarem que té un comportament lineal. Per al NIHSS basal, el canvi d'octil té una tendència lineal. Així, en el model la introduïrem categoritzada. Per a l'edat també tenim una tendència lineal per al canvi d'octil. També, la introduïrem categoritzada dintre del model.

La taula 15 presenta els cumulative OR ajustats i no ajustat per a les variables categòriques, la temperatura com a quantitativa i el NIHSS basal i l'edat categoritzada en octils. Així, l' OR de l'augment d'un grau de la temperatura basal sobre les seqüeles és de 1,0379 no ajustat i de 0,8864 ajustat. Aquests OR no són significativament diferents de 1. Per tant, la temperatura basal no ens serveix com a variable pronòstic de les seqüeles a les 12 setmanes.

La citicolina té un cumulative OR de gravetat de les seqüeles no ajustat de 0,9132 i ajustat de 0,8161 (taula 15); però l'efecte protector de la citicolina és no significatiu. Per tant, el tractament no està associat a tenir unes seqüeles menys greus. L'efecte ajustat o no del sexe sobre les seqüeles són no significatius. A diferència del vist fins ara, la raça negra té un pitjor odds ajustat que la blanca.

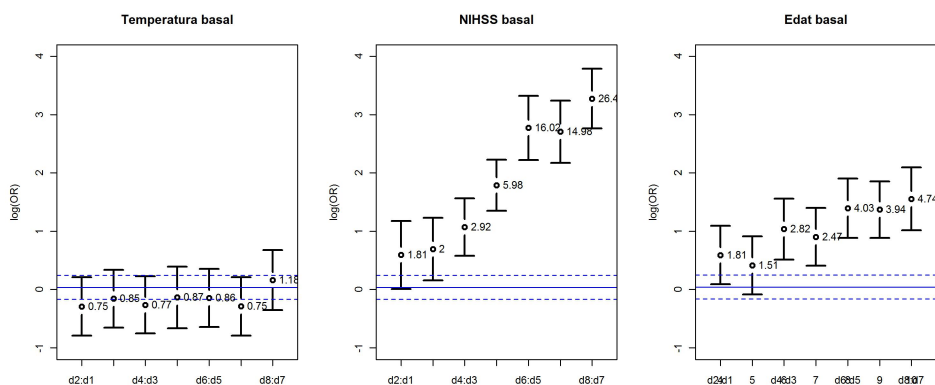


FIGURA 19. Foresplot del $\ln(\text{cumulative } OR)$ del canvi d'octil per a cada variable quantitativa. Les línies blaves marquen el valor del $\ln(\text{cumulative } OR)$ sota el supòsit de linealitat i el seu interval de confiança al 95%. Els valors mostrats són els diversos cumulative OR de cada octil.

Les diverses categories del NIHSS tenen un OR significativament diferents de 1. Com el $NIHSS$ basal sigui més alt, més odds de tenir unes seqüeles més greus. A més, el creixement dels OR no és lineal. Així, el canvi en les categories baixes no és el mateix que en les categories altes. Els diversos octils de l'edat tenen un OR significativament diferents de 1. Com més edat tingui el pacient, té més odds de patir seqüeles greus.

Per tant, si el model és complet i correcte, concloem que l'augment d'un grau de la temperatura basal no suposa cap efecte sobre l'odds de patir seqüeles més greus. Tampoc el tractament suposa una reducció de l'odds de tenir seqüeles severes respecte el placebo. En canvi, l'augment del NIHSS basal i de l'edat comporten un pitjor pronòstic de patir pitjors seqüeles a les 12 setmanes.

Validació del model Els estadístics de la bondat de l'ajust i de la capacitat predictiva estan a la taula 16. El test de la raó de la versemblança té un p -valor $< 0,0001$. Així, rebutjem la hipòtesis de que el model complet sigui equivalent al model buit.

Si ens fixem en la capacitat predictiva, l'àrea sota la corba ROC indica que el model discrimina de manera acceptable. Però els diversos tests de Hosmer-Lemeshow tenen uns p -valors molt baixos, indicació de que els diferents models no són del tot correctes.

Per tant, conclouem que el model construït discrimina de manera acceptable; però no és correcte. Tot i que amb les variables d'estudi no podem construir-ne cap de millor. A part, hem comprovat que simplificar els diversos OR en un sol cumulative OR és coherent i que el supòsit de linealitat és acceptable per a la temperatura basal.

Variable	Categoria	OR	CI(OR, 95%)	OR	CI(OR, 95%)
		no ajustat		ajustat	
Temperatura basal	1 unitat	1,0379	[0,8457; 1,2739]	0,8864	[0,7116; 1,1043]
Tractament	Placebo	Ref		Ref	
	Citicolina	0,9132	[0,7125; 1,1705]	0,8519	[0,6460; 1,1235]
Sexe	Dona	Ref		Ref	
	Home	0,8180	[0,6380; 1,0489]	1,0864	[0,8200; 1,4393]
Races	Blanca	Ref		Ref	
	Negra	1,1750	[0,8276; 1,6682]	1,5899	[1,0802; 2,3403]
	Altres	1,1202	[0,6555; 1,9142]	0,8805	[0,4869; 1,5925]
NIHSS	$7 < NIHSS \leq 9$	Ref		Ref	
	$9 < NIHSS \leq 10$	1,8100	[1,0099; 3,2438]	1,7978	[0,9635; 3,3546]
	$10 < NIHSS \leq 11$	2,0015	[1,1701; 3,4235]	2,0213	[1,1386; 3,5883]
	$11 < NIHSS \leq 13$	2,9217	[1,7843; 4,7841]	2,9435	[1,7398; 4,9800]
	$13 < NIHSS \leq 16$	5,9823	[3,8609; 9,2692]	7,0837	[4,4456; 11,2874]
	$16 < NIHSS \leq 17$	16,0182	[9,2399; 27,7689]	17,8764	[10,0164; 31,9041]
	$17 < NIHSS \leq 20$	14,9847	[8,7770; 25,5829]	16,5370	[9,3123; 29,3667]
Edat	$20 < NIHSS \leq 34$	26,4365	[15,8241; 44,1660]	29,8270	[17,1477; 51,8817]
	$19 < Edat \leq 52$	Ref		Ref	
	$52 < Edat \leq 60$	1,8069	[1,0938; 2,9849]	1,9081	[1,0711; 3,3995]
	$60 < Edat \leq 66$	1,5135	[0,9197; 2,4908]	2,0773	[1,1755; 3,6709]
	$66 < Edat \leq 70$	2,8199	[1,6714; 4,7574]	2,9635	[1,6435; 5,3437]
	$70 < Edat \leq 74$	2,4665	[1,5034; 4,0466]	2,6462	[1,5126; 4,6294]
	$74 < Edat \leq 77$	4,0317	[2,4166; 6,7261]	4,9969	[2,7832; 8,9714]
	$77 < Edat \leq 80.5$	3,9385	[2,4262; 6,3935]	4,4922	[2,5747; 7,8378]
$80.5 < Edat \leq 85$	4,7357	[2,7572; 8,1339]	6,6171	[3,5498; 12,3347]	

TAULA 15. L'odds ratio de tenir pitjors seqüeles no ajustat i ajustat pel canvi de categoria per les variables categòriques i per a l'augment d'una unitat en les numèriques.

Estadístic de validació	Valor	p-valor
Raó de versemblança	320,68	$\leq 0,0001$
D de Sommers	0,562	
Gamma	0,565	
Tau de Kendall	0,376	
Capacitat predictiva	Valor	p-valor
Àrea sota la corba ROC	0,781	
Test de Hosmer-Lemeshow: $y \geq 3$		$\leq 0,0001$
Test de Hosmer-Lemeshow: $y \geq 5$		0,7717
Test de Hosmer-Lemeshow: $y \geq 6$		$\leq 0,0001$

TAULA 16. Estadístics de validació del model logístic amb odds proporcionals per a les seqüeles a les 12 setmanes.

Capítol 4

Discussió

1. Resum dels resultats

El nostre estudi buscava una relació entre l'evolució de la temperatura i el tractament rebut. Però no s'ha aconseguit demostrar que existeixi cap relació. En canvi, sí que s'ha pogut explicar l'evolució de la temperatura respecte el temps de seguiment. El segon objectiu era quantificar el risc de morir segons la temperatura basal. No s'ha aconseguit demostrar que la temperatura basal influeix en la mort del pacient. La última qüestió tractava sobre l'efecte de la temperatura basal sobre les seqüeles a les dotze setmanes. Tampoc s'ha aconseguit provar que l'efecte de la temperatura basal influeixi en el valor de l'escala *mRS* a les dotze setmanes.

En canvi, sí que s'ha pogut establir que les seqüeles basals mesurades amb l'escala NIHSS estan relacionades amb la mortalitat i les seqüeles als 3 mesos mesurades amb l'escala *mRS*. S'ha observat que l'evolució del homes i de les dones es diferent, ja que existeix una interacció entre el sexe i el terme lineal del temps de seguiment ($S*t = -0,001633$). També s'ha observat que l'augment del NIHSS basal suposa un augment de l'odds de morir en 19% ajustat per les altres variables d'ajust. També suposa un augment de l'odds de tenir seqüeles més severes.

2. Models estadístics

Si ens centrem en la part estadística, hem estudiat un problema real del qual teníem unes hipòtesis que volíem respondre. L'estudi s'ha centrat en estudiar les tres hipòtesis per respondre-les de la millor manera possible. Per modelar l'evolució de la temperatura, hem elegit un model lineal mixt, perquè l'evolució de la temperatura té una resposta contínua i a cada visita es distribueix com una normal. Per quantificar l'efecte sobre la mortalitat o les seqüeles, s'ha utilitzat el model logístic. Gràcies als *OR* que es poden calcular amb el model, s'ha pogut establir el risc de morir o de patir seqüeles respecte les variables d'estudi.

Per a l'evolució de la temperatura corporal, hem construït el millor model amb les variables d'estudi. Hem comprovat que el model construït era raonable. Després

s'ha provat l'efecte de la citicolina sobre l'evolució temporal de la temperatura. No s'ha pogut demostrar que la citicolina modifiqués l'evolució de la temperatura.

Els dos models logístics que s'han exposat no són capaços de recollir la major part de variabilitat del model. Aquest fet pot ser degut a varies causes com la falta d'un factor explicatiu important, una mala modelització, etc.

3. Explicació de l'efecte observat

Quan un pacient pateix una AVC isquèmic, inicia un procés de degeneració de les cèl·lules nervioses. Molts cops aquest procés acabarà sent irreversible. Un AVC isquèmic està produïda per una reducció dràstica del flux sanguini cerebral. Totes les intervencions sobre el pacient van destinades a restablir el flux normal de sang.

En l'assaig clínic estudiat [3], no es va demostrar que la citicolina sigui un tractament eficaç per a reduir ni les seqüeles als tres mesos ni la temperatura corporal.

Com que no existeix cap intervenció de restabliment del flux, l'esforç va dirigit a guanyar temps i reduir les possibles seqüeles. Una de les possibles intervencions per a reduir les seqüeles és disminuir la temperatura corporal del pacient. Amb aquesta intervenció, es vol aconseguir que les cèl·lules nervioses metabolitzin més lentament i el deteriorament sigui més lent.

De l'explicació prèvia, si una cèl·lula té una temperatura alta metabolitzarà més ràpid, és més probable que el pacient tingui pitjors seqüeles o morir. Per tant, pensàvem que els pacients amb una temperatura basal alta tindrien una major mortalitat i unes pitjors seqüeles. Però no s'observa aquest fenomen, perquè la temperatura corporal basal només mesura la temperatura al començament del tractament. Necessitaríem una mesura de l'evolució durant les primeres hores del tractament, ja que és quan es produeixen les seqüeles més importants.

4. Comparació amb altres articles

Hi han molts articles que han estudiat la relació entre la temperatura corporal i els AVC isquèmics. En els diferents estudis, la temperatura corporal juga diversos papers, ja sigui com a variable principal (p.e. [16]), com a variable d'ajust (p.e. [13]) o com a variable resposta (p.e. [8]).

Només hem trobat un assaig clínic centrat en la intervenció d'un tractament sobre la temperatura [8]. L'article de Dippel [8] és un assaig clínic amb tres branques (placebo, acetaminofé i ibuprofé) on suposen que els medicaments produeixin una disminució de la temperatura durant les primeres 24 hores. Però no troben diferències significatives entre les tres branques. Així, que tenen resultats semblants als nostres, ja que tampoc poden confirmar que les intervencions facin disminuir la temperatura corporal.

Hi han variis assaig clínics que han comprovat si la hipotermia induïda era un bon tractament. La hipotermia va ser produïda per un medicament o una intervenció física (p.e. bosses de gel).

L'article de Holzer [11] està centrat en les seqüeles severes i la supervivència a 6 mesos. Als 6 mesos, hi ha un 55% (75/136) dels pacients hipotèrmics que tenien un índex de Pittsburgh igual o inferior a 2. En canvi, en el grup control només el 39% (54/137) tenien un índex igual o inferior a 2. Ajustat per l'edat, el sexe i altres variables clíniques, van trobar que el risc relatiu de sobreviure era de 0,62 (95%CI: 0,36-0,95) pels hipotèrmics enfront dels normotèrmics. La supervivència als 6 mesos és un 14% superior en el grup intervingut. Els resultats d'aquests article milloren els resultats que hem trobat de supervivència, ja que nosaltres no hem trobat cap relació entre la temperatura i la mort.

Els assaig clínics de Bernard [1] i Marion [15] estudien l'efecte de la hipotermia sobre les seqüeles. L'article de Bernard [1] està centrat en les altes clíniques. En canvi, l'article de Marion [15] estudia les seqüeles a mitjà termini, fins als 12 mesos. L'assaig de Bernard [1] calcula que la hipotermia millora l'odds de les seqüeles respecte la normoterapia ($OR = 5,25$ 95%CI : 1,47 – 18,76) ajustat per l'edat, el sexe i altres variables clíniques. En canvi, l'assaig de Marion [15] no pot concloure que la hipotermia millori la supervivència als 3, 6 o 12 mesos, perquè els intervals de confiança pels OR ajustats contenen el 1.

El meta-anàlisis de Harris [10] estudia 7 assaigs clínics en que la intervenció sobre diferents danys cerebrals era la hipotermia. Els diferents assaigs estudiats tenen diversos tamanyos mostrals i la inducció de la hipotermia es fa de manera diferent. Calculen que la normoterapia té un odds pitjor que la hipotermia, $OR = 0,61$ (95%CI : 0,26 – 1,46). Aquest OR es calcula amb el mètode de DerSimonian i Laird, ja que les dades presenten heterogeneïtat. Arriben a la mateixa conclusió que nosaltres, ja que no poden demostrar que les temperatures baixes millorin les seqüeles.

L'estudi de Kammergaard [13] és prospectiu i no avalua cap intervenció. Estudien la supervivència fins als 5 anys segons si el pacient tenia una temperatura superior als $37^{\circ}C$. S'observa que la disminució d'un grau en la temperatura té associat un hazard ratio (HR) a 5 anys de 1,30 (95%CI : 1,04 – 1,57) ajustat pel SSS inicial, l'edat i altres variables.

També, hi han diversos articles epidemiològics que han estudiat les seqüeles de les malalties cerebrovasculars isquèmiques. Aquests volen estudiar la relació causal entre la temperatura i les seqüeles.

L'article de den Hertog [6] utilitza les dades de l'estudi PAIS (PARacetamol In Stroke). Estudien la relació entre la temperatura basal o l'evolució durant les primeres 24 hores amb la severitat de les seqüeles i la mort del pacient. Observen que la temperatura basal no té cap efecte sobre les seqüeles o la mort. En canvi, si la temperatura augmenta durant les primeres 24 hores, l'odds de patir seqüeles augmenta un 30% ($OR = 1,30$ 95%CI : 1,05 – 1,63). Aquest OR ha estat ajustat per l'edat, el sexe, la temperatura basal, el NIHSS basal i altres variables. Només

podem comparar els resultats sobre la temperatura basal els quals són similars als nostres.

L'article de Millan [17] és un estudi retrospectiu amb 254 pacients. Van observar que els pacients amb una temperatura superior als 37°C tenien associat un $OR = 2,56$ ($95\%CI : 1,19 - 5,50$) de patir seqüeles enfront de la resta de pacients. Aquest OR està ajustat per l'edat, el NIHSS basal i altres variables mèdiques. El seu resultat mostra una relació més forta que la nostra, però també inclouen més malalties cerebrovasculares.

L'article de Reith [20] utilitza les dades de l'estudi prospectiu de Copenhagen amb 1197 pacients. Analitzen la severitat i tamany de la malaltia cerebrovascular i les seqüeles amb tres models lineals diferents. També estudien la supervivència i les seqüeles a l'alta mèdica amb dos models logístics. Van observar que si la temperatura augmenta un 1°C l' OR de morir era de 1,8 ($95\%CI : 1,1 - 2,8$) ajustat per la severitat inicial i si el pacient havia tingut un AVC previ. Per a les seqüeles, un augment de 1°C té associat un OR de tenir pitjors seqüeles de 2,2 ($95\%CI : 1,4 - 3,5$) ajustat només per la severitat inicial. L'efecte mesurat és més gran que el nostre, però és possible que l'efecte observat es redueixi en el temps. Ja que estudien la supervivència a curt termini (pocs dies) i nosaltres hem estudiat a mitjà termini (3 mesos).

Els diversos articles estudiats arriben a diverses conclusions; però en cada un la temperatura té un rol diferent i hem de ser cauts a l'hora de comparar els seus resultats amb els nostres. L'únic article que la temperatura és la seva variable intervinguda obté resultats semblants als nostres, però només monotoritzen les primeres 24 hores. Si ens centrem en la mortalitat o les seqüeles, tenim nombrosos articles. Pocs conclouen que l'augment de la temperatura basal comporta un augment de morir o de tenir pitjors seqüeles. En canvi, els articles que utilitzen la variació durant les primeres hores sí que apunten que una evolució que augmenta comporta un pitjor pronòstic.

5. Limitacions

Aquest estudi presenta diverses limitacions que no s'han pogut corregir. El més important és que les dades van ser recollides per mesurar una intervenció específica i la temperatura no era la seva variable principal. Així, la temperatura no va ser mesurada de manera sistemàtica i seguint un protocol. Aquesta font de variabilitat és molt gran, perquè la temperatura té una variabilitat que no podem controlar o quantificar i, probablement, depèn de cada hospital. Un altre tipus de limitació és el biaix de les pèrdues, ja que fa perdre l'aleatoreïtat de les dades i pot estar sota-estimant l'efecte sobre de la variable d'estudi.

6. Recerca futura

Finalment, hem conclòs que la citicolina no millora la supervivència o redueix les seqüeles en els pacients que han patit un AVC isquèmic. Hi han molts articles i

estudis que han analitzat diverses intervencions o la relació de la temperatura amb la supervivència o les seqüeles. Però de tots els articles analitzats s'han trobat resultats molt diferents. Per tant, s'ha de seguir estudiant les causes biològiques de l'AVC isquèmic per a trobar alguna intervenció eficient i comprovar-la en futurs assaigs clínics.

Apèndix A

Construcció del model sense tractament

En aquest annex explicarem la construcció del model sense tractament amb tots els detalls.

El primer model que s'ha provat és una aproximació lineal de la temperatura segons només el dia de seguiment, $T(t) = C_0 + C_1 \cdot t$. La sortida de R per aquest model és:

Coefficients:

	Estimate	Std. Error	t value	$Pr(> t)$
(Intercept)	36.678966	0.011853	3094.536	< 2e-16 ***
VS.DAY	-0.002291	0.000299	-7.663	2.25e-14 ***

La figura 1 presenta els residus del model enfront dels valors estimats. El model lineal és rígid i no s'ajusta a les dades de l'estudi, ja que per a cada valor del temps de seguiment només tenim un sol valor predit. Necessitem flexibilitzar l'ajust, ja que totes les mesures no presenten una tendència lineal global.

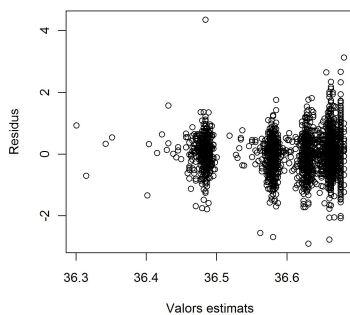


FIGURA 1. Bivariat entre els residus estandaritzats i les dades reals per al model lineal sobre la temperatura.

Flexibilitarem el model permeten que els paràmetres del model no siguin iguals per a tots els pacients. Així, la introducció d'un efecte aleatori ens permetrà ajustar millor la temperatura amb una estimació diferent per a cada pacient del mateix model. El primer model que s'ha contrastat és un model lineal respecte el dia de seguiment on el pendent té associat un efecte aleatori. Així, el model és

$$T_p(t) = C_{p,0} + C_{p,1} \cdot t$$

on $C_{p,0} \sim N(C_0, \sigma_{C_0}^2)$ i $C_{p,1} \sim N(C_1, \sigma_{C_1}^2)$. Cada pacient té la seva pròpia estimació. L'estimació del model és:

Random effects:

Formula: \sim VS.DAY | PATID

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.258670333	(Intr)
VS.DAY	0.003194849	-0.671
Residual	0.506844751	

Fixed effects: TEMPC \sim VS.DAY

	Value	Std.Error	DF	t-value	p-value
(Intercept)	36.68086	0.013980168	3194	2623.7785	0
VS.DAY	-0.00228	0.000298887	3194	-7.6131	0

Correlation:

	(Intr)
VS.DAY	-0.639

La figura 2 presenta el diagrama bivariat entre els residus estandaritzats i les dades reals. Observem que existeix una tendència dintre dels residus. Aquesta tendència, destaca que les estimacions dels valors baixos i dels valors alts es queden curtes. Per tant, el model no és capaç de captar tota la variabilitat de les dades.

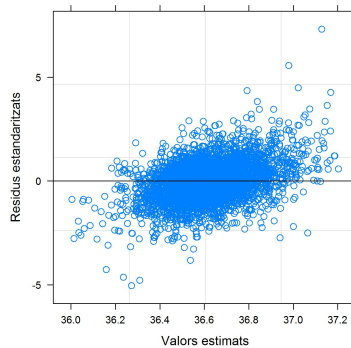


FIGURA 2. Bivariat entre els residus estandaritzats i les dades reals per al model lineal amb efectes aleatoris sobre la temperatura.

Com que el model lineal, no s'ajusta correctament, introduïrem el terme quadràtic del temps de seguiment. A més, els tres paràmetres tenen associat un efecte aleatori.

El model és

$$T(t) = C_{p,0} + C_{p,1} \cdot t + C_{p,2} \cdot t^2$$

on $C_{p,0} \sim N(C_0, \sigma_{C_0}^2)$, $C_{p,1} \sim N(C_1, \sigma_{C_1}^2)$ i $C_{p,2} \sim N(C_2, \sigma_{C_2}^2)$. El model quadràtic té problemes de convergència numèrica. Si incloem la funció de variància potència de la co-variable més una constant, sí que convergeix. La estimació d'aquest model és

Random effects:

```
Formula: ~ VS.DAY + I(VS.DAY^2) | PATID
Structure: General positive-definite, Log-Cholesky parametrization
          StdDev          Corr
(Intercept) 0.2819618593 (Intr) VS.DAY
          VS.DAY 0.0130426406 -0.611
I(VS.DAY^2) 0.0001110662 0.534 -0.971
Residual 0.5811449621
```

Variance function:

```
Structure: Power of variance covariate
Formula: ~VS.DAY
Parameter estimates:
```

```
power
-0.07558437
```

Fixed effects: TEMPC ~ VS.DAY + I(VS.DAY^2)

	Value	Std.Error	DF	t-value	p-value
(Intercept)	36.72450	0.017381041	3193	2112.9054	0
VS.DAY	-0.00734	0.001052042	3193	-6.9811	0
I(VS.DAY^2)	0.00006	0.000010876	3193	5.3776	0

Correlation:

	(Intr)	VS.DAY
VS.DAY	-0.724	
I(VS.DAY^2)	0.610	-0.961

Compararem els dos models amb el test de la raó de versemblances. La millora del model quadràtic és substancial com es pot veure en la següent taula:

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
lineal	1	6684.427	6722.323	-3336.213			
Quadràtic	2	6574.813	6644.290	-3276.407	1 vs 2	119.6135	< .0001

La figura 3 presenta el bivariat entre les estimacions i els residus estandaritzats. Amb el terme quadràtic no s'ha eliminat la tendència entre els residus i les estimacions. Així, l'estimació no és bona.

La introducció de més termes polinòmics per al temps de seguiment o de les variables d'ajust no portarà a millors ajustos, ja que les aproximacions no són capaços d'ajustar-se a la forma de la temperatura. Necessitem aplicar una transformació de les dades per a trobar un millor ajust. Pensem que una transformació de Box-Cox no és necessària, ja que el possible problema de no homocedasticitat el podem solucionar amb una funció de variància. Tampoc pensem que aplicar una diferència

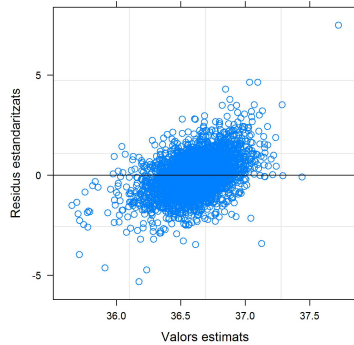


FIGURA 3. Bivariat entre els residus estandaritzats i les dades reals per al model quadràtic amb efectes aleatoris i funció de variància potència d'una covariable més una constant sobre la temperatura.

de primer ordre a les dades sigui necessària, perquè les dades ja són estacionàries al voltant de 36.5°C . Pensem que la transformació ha de ser la diferència respecte a la basal, $dT(t) = T(t) - T(0)$. Ja que ens permetrà estudiar la forma de l'evolució sense tenir en compte el nivell a que està. A més, totes les corbes partiran del mateix valor.

El primer model estimat és el model lineal sobre la diferència respecte la basal. El model estimat és

Random effects:

Formula: $\sim \text{VS.DAY} \mid \text{PATID}$

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr
(Intercept)	0.615693633	(Intr)
VS.DAY	0.003573305	-0.15
Residual	0.466308819	

Fixed effects: $I(\text{TEMPC} - \text{TEMPB}) \sim \text{VS.DAY}$

	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.07063827	0.025637339	2294	-2.755289	0.0059
VS.DAY	-0.00156056	0.000331728	2294	-4.704333	0.0000

Correlation:

	(Intr)
VS.DAY	-0.397

La figura 4 presenta el bivariat entre les estimacions i els residus estandaritzats. La tendència que observàvem abans, ara és menys accentuada. A més, només el 1.88% dels residus estan fora de les bandes $[-1, 96; 1, 96]$. Així, la millora d'aquest ajust és significativa. A part, no podem comparar aquest model amb els anteriors, ja que els conjunts de dades reals no són el mateix.

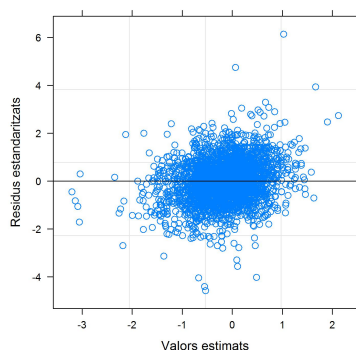


FIGURA 4. Bivariat entre els residus estandaritzats i les dades reals en el model lineal amb efectes aleatoris sobre la diferència respecte la basal.

En el següent model incloem el terme quadràtic del temps de seguiment amb efecte aleatori. Pensem que potser correcta, ja que l'evolució del pacient no té perquè ser lineal. Poden tenir una disminució més accentuada al principi que al final de l'estudi. El terme quadràtic ens permetrà capturar aquesta tendència, però hem de tenir en compte que l'ajust només serà vàlid dintre del període d'estudi. El model estimat és

Random effects:

Formula: $\sim VS.DAY + I(VS.DAY^2) \mid PATID$

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr		
(Intercept)	0.6587052072		(Intr)	VS.DAY
VS.DAY	0.0171444191	-0.383		
I(VS.DAY ²)	0.0001359866	0.378	-0.998	
Residual	0.4450014262			

Fixed effects: $I(TEMPC - TEMPB) \sim VS.DAY + I(VS.DAY^2)$

	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.006304127	0.030334549	2293	-0.207820	0.8354
VS.DAY	-0.006569391	0.001336982	2293	-4.913595	0.0000
I(VS.DAY ²)	0.000052085	0.000013116	2293	3.971026	0.0001

Correlation:

	(Intr)	VS.DAY
VS.DAY	-0.611	
I(VS.DAY ²)	0.549	-0.969

La següent taula presenta diferents versions dels models lineals i quadràtics presentats. S'han inclòs diverses funcions de variàncies.

Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
lineal +varPower	1 7	5745.192	5787.423	-2865.596			
lineal	2 6	5770.452	5806.650	-2879.226	1 vs 2	27.26034	< .0001
lineal +varExp	3 7	5737.786	5780.017	-2861.893	2 vs 3	34.66631	< .0001
quadràtic	4 10	5744.155	5804.482	-2862.078			
quadràtic + varConstPower	5 12	5737.089	5809.481	-2856.544	4 vs 5	11.06647	0.004
lineal + varExp	3 7	5717.881	5760.117	-2851.941			
lineal + varConstPower	5 12	5696.513	5768.917	-2836.256	1 vs 2	31.36864	< .0001

Concloem que el millor model és el quadràtic amb efectes aleatoris quadràtics i una funció de variància que és la potència del temps de seguiment més una constant.

La figura 5 presenta el bivariat entre les estimacions i els residus estandaritzats. Amb el terme quadràtic no s'ha eliminat la tendència entre els residus i les estimacions. Així, l'estimació no és bona.

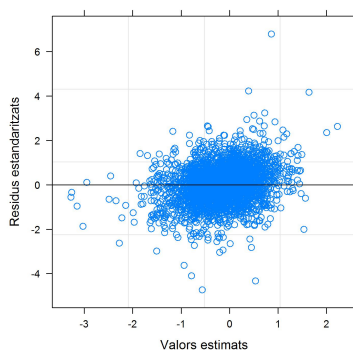


FIGURA 5. Bivariat entre els residus estandaritzats i les dades reals per al model quadràtic amb efectes aleatoris i funció de variància potència d'una covariable més una constant sobre la diferència de temperatura respecte la basal.

La introducció de nous termes polinòmics del temps de seguiment suposarà un sobreajust del model. El pròxim pas per a millorar les estimacions és la introducció de les variables d'estudi. Introduïrem cada vegada una variable d'estudi diferent. A més, suposem que les variables d'estudi tenen un efecte igual per a tots els pacients. Les variables entraran el model modifiquen la constant, el terme lineal i el quadràtic del temps de seguiment.

La primera variable que introduïm és el sexe del pacient. Només mostrarem el test de raó de versemblances per als diferents models possibles. El primer model conté els tres paràmetres associats al sexe. En el segon, s'ha eliminat el terme quadràtic per no ser significatiu. En el tercer, només té el terme lineal del sexe. En la taula següent tenim els diversos tests comparant un model respecte el sobre:

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
model + SEX + SEX:t								
+ SEX:t ²	1	15	5694.29	5784.79	-2832.14			
model + SEX + SEX:t	2	14	5692.35	5776.82	-2832.18	1 vs 2	0.0657	0.7977
model + SEX:t	3	13	5691.04	5769.48	-2832.52	2 vs 3	0.6853	0.4078
Model quadràtic RE i varConstPower	4	12	5696.51	5768.92	-2836.26	3 vs 4	7.4742	0.0063

El model que incorpora el terme lineal del sexe ajusta millor que el quadràtic sol. Així, introduïm aquest terme al model.

La segona variable introduïda és la raça. En el primer model s'ha incorporat els tres termes possibles de la raça. En el segon només tenim el que modifica la constant i el terme quadràtic. El tercer model només conté el paràmetre que modifica la constant segons la raça. La taula per al test de raons de versemblança és:

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
model + SEX*t + RACE +								
RACE:t + RACE:t ²	1	19	5698.23	5812.87	-2830.12			
model + SEX*t + RACE +	2	17	5695.40	5797.97	-2830.70	1 vs 2	1.1703	0.5570
RACE:t ²								
model + SEX*t + RACE	3	15	5692.69	5783.20	-2831.35	2 vs 3	1.2929	0.5239
Model + SEX*t	4	13	5691.04	5769.48	-2832.52	3 vs 4	2.3457	0.3095

En aquest cas, la raça no entra a formar part del model, perquè cap dels models possibles millora l'ajust del model amb el sexe.

La tercera variable d'ajust és el NIHSS basal. També com en els dos models previs, introduïm la variable juntament amb la interacció amb el temps de seguiment i amb el terme quadràtic del seguiment. La taula de raó de versemblances és:

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
Model + SEX*t + BNIH+								
BNIH*t+BNIH*t ²	1	16	5688.99	5785.52	-2828.49			
Model + SEX*t +								
BNIH*t+BNIH*t ²	2	15	5687.28	5777.78	-2828.64	1 vs 2	0.2922	0.5888
Model + SEX*t + BNIH*t ²	3	14	5692.00	5776.47	-2832.00	2 vs 3	6.7251	0.0095
Model + SEX*t	4	13	5691.04	5769.48	-2832.52	3 vs 4	1.0364	0.3087
						2 vs 4	7.7615	0.0206

El model que inclou el sexe i les interaccions entre el NIHSS basal amb el terme lineal i el terme quadràtic millora l'ajust del model incloent el sexe. El problema rau en que la interacció amb el terme lineal és no significativa i es tendeix a la sobre-parametrització del model, com el criteri *BIC* indica.

La última variable que introduïm és l'edat. Com en els models precedents, el primer model conté l'edat i les dues interaccions. En el segon s'elimina la interacció amb el terme quadràtic. El tercer model només conté l'edat. La taula amb el test de raó de versemblances és:

	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
Model + SEX*t + AGE + AGE*t + AGE* ²	1	16	5694.55	5791.09	-2831.28			
Model +SEX*t + AGE + AGE*t	2	15	5692.59	5783.09	-2831.29	1 vs 2	0.0349	0.8519
Model +SEX*t + AGE	3	14	5691.92	5776.39	-2831.96	2 vs 3	1.3305	0.2487
Model + SEX*t	4	13	5691.04	5769.48	-2832.52	3 vs 4	1.1233	0.2892

L'edat tampoc millora l'ajust del model per al temps de seguiment juntament amb la interacció entre el sexe i el terme lineal del seguiment.

El model que ajusta millor les dades és un model quadràtic respecte el temps de seguiment amb una interacció entre el sexe i el terme lineal del seguiment. El model conté uns efectes aleatoris sobre la constant, el terme lineal del seguiment i el terme quadràtic del seguiment. A més, la variància dels residus intra-individus disminueixen segons la potència del seguiment més una constant. La validació i estimació del model està en el cos del text.

1. Model amb pacients amb totes les mesures

Per assegurar-nos que l'efecte del tractament és no significatiu, estimarem el model sobre els pacients amb totes les mesures. Ja que les mesures perdudes podrien estar esbiaixan el resultat.

Primer estimem el model sense el tractament:

Random effects:

Formula: \sim VS.DAY + I(VS.DAY²) | PATID

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr		
(Intercept)	0.6316834748		(Intr)	VS.DAY
VS.DAY	0.0177376586	-0.263		
I(VS.DAY ²)	0.0001412673	0.218	-0.990	
Residual	0.2316289503			

Variance function:

Structure: Constant plus power of variance covariate

Formula: \sim VS.DAY

Parameter estimates:

const	power
1.04444753	-0.02418353

Fixed effects: I(TEMPC - TEMPB) \sim VS.DAY + I(VS.DAY²) + SEX:VS.DAY

	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.06258336	0.04709603	1153	-1.328846	0.1842
VS.DAY	-0.00493220	0.00209715	1153	-2.351855	0.0188
I(VS.DAY ²)	0.00004759	0.00001996	1153	2.384835	0.0172
VS.DAY:SEXMale	-0.00217136	0.00091989	1153	-2.360460	0.0184

Correlation:

	(Intr)	VS.DAY	I(VS.D
VS.DAY	-0.545		
I(VS.DAY ²)	0.496	-0.949	
VS.DAY:SEXMale	0.003	-0.220	0.011

El model amb tractament és

Random effects:

Formula: \sim VS.DAY + I(VS.DAY²) | PATID

Structure: General positive-definite, Log-Cholesky parametrization

	StdDev	Corr	
(Intercept)	0.6320921863	(Intr)	VS.DAY
VS.DAY	0.0177528738	-0.269	
I(VS.DAY ²)	0.0001412135	0.225	-0.991
Residual	0.2233134395		

Variance function:

Structure: Constant plus power of variance covariate

Formula: \sim VS.DAY

Parameter estimates:

const	power
1.11098215	-0.02133115

Fixed effects: I(TEMPC - TEMPB) \sim VS.DAY + I(VS.DAY²)+VS.DAY:SEX +VS.DAY:TREAT + I(VS.DAY²):TREAT

	Value	Std.Error	DF	t-value	p-value
(Intercept)	-0.06167074	0.04711371	1151	-1.308976	0.1908
VS.DAY	-0.00755075	0.00279856	1151	-2.698086	0.0071
I(VS.DAY ²)	0.00007116	0.00002762	1151	2.576329	0.0101
VS.DAY:SEXMale	-0.00213472	0.00092024	1151	-2.319736	0.0205
VS.DAY:TREATCiticoline	0.00485665	0.00340163	1151	1.427742	0.1536
I(VS.DAY ²):TREATCiticoline	-0.00004375	0.00003476	1151	-1.258563	0.2084

Correlation:

	(Intr)	VS.DAY	I(VS ²)	VS.DAY:S	VS.DAY:T
VS.DAY	-0.422				
I(VS.DAY ²)	0.375	-0.955			
VS.DAY:SEXMale	0.003	-0.179	0.017		
VS.DAY:TREATCiticoline	0.018	-0.662	0.660	0.021	
I(VS.DAY ²):TREATCiticoline	-0.021	0.642	-0.691	-0.014	-0.963

En els pacients que van venir a totes les visites, tampoc es demostra que l'efecte de la Citicolina sigui significatiu.

Apèndix B

Taules de la linealitat dels logits de morir

1. Taules dels logits

Les taules següents contenen el nombre de vius i morts que tenim a cada quartil. A més, hi ha l'*OR* de morir quan es canvia del decil de la i al decil $i + 1$, juntament amb el seu interval de confiança al 95%.

Decil	Vius		Morts		OR	IC(OR,95%)
	i	i+1	i	i+1		
34,60 - 36,00	74	65	23	12	0.5940	0,2741 - 1,2872
36,00 - 36,22	65	75	12	8	0.5778	0,2225 - 1,5004
36,22 - 36,44	75	81	8	11	1.2732	0,4858 - 3,3363
36,44 - 36,63	81	82	11	15	1.3470	0,5836 - 3,1090
36,63 - 36,80	82	66	15	14	1.1596	0,5225 - 2,5736
36,80 - 37,00	66	60	14	12	0.9429	0,4044 - 2,1985
37,00 - 37,17	60	78	12	14	0.8974	0,3869 - 2,0815
37,17 - 37,44	78	69	14	15	1.2112	0,5458 - 2,6878
37,44 - 39,80	69	58	15	18	1.4276	0,6616 - 3,0803

TAULA 1. Taula dels *OR* de morir al canviar el decil de la temperatura basal.

Decil	Vius		Morts		OR	IC(OR,95%)
	i	i+1	i	i+1		
7-8	102	75	0	4		
8-10	75	145	4	13	1,6810	0,5297 - 5,3347
10-11	145	51	13	8	1,7496	0,6857 - 4,4645
11-13	51	48	8	7	0,9297	0,3131 - 2,7604
13-14	48	94	7	16	1,1672	0,4497 - 3,0296
14-16	94	56	16	16	1,6786	0,7788 - 3,6177
16-17	56	76	16	35	1,6118	0,8126 - 3,1970
17-21	76	51	35	20	0,8515	0,4428 - 1,6377
21-34	51	48	20	34	1,8063	0,9164 - 3,5601

TAULA 2. Taula dels *OR* de morir al canviar el decil del *NIHSS* basal.

Decil	Vius		Morts		OR	IC(OR,95%)
	i	i+1	i	i+1		
19-52	92	80	6	6	1,1500	0,3567 - 3,7075
52-58	80	72	6	15	2,7778	1,0230 - 7,5423
58-64	72	89	15	13	0,7011	0,3134 - 1,5683
64-69	89	81	13	17	1,4369	0,6571 - 3,1418
69-72	81	83	17	19	1,0907	0,5296 - 2,2462
72-75	83	54	19	17	1,3752	0,6570 - 2,8785
75-78	54	60	17	21	1,1118	0,5317 - 2,3246
78-81	60	73	21	12	0,4697	0,2138 - 1,0320
81-85	73	62	12	27	2,6492	1,2394 - 5,6625

TAULA 3. Taula dels *OR* de morir al canviar el decil de l'edat.

2. Assumpció de linealitat dels logits de l'edat

Com que la suposició de linealitat de l'edat no queda del tot clara, hem estudiat el logaritmes de odds per a diverses categoritzacions de l'edat, taula 4.

Decil	Vius	Morts	log(odds)	$IC_{95\%}$
19-51	92	6	2,7300	1,90418 - 3,5559
51-57	80	6	2,5903	1,76064 - 3,4199
57-62	72	15	1,5686	1,01232 - 2,1249
62-67	89	13	1,9237	1,34173 - 2,5056
67-71	81	17	1,5612	1,03836 - 2,0841
71-74	83	19	1,4744	0,97593 - 1,9729
74-76	54	17	1,1558	0,61069 - 1,7009
76-78	60	21	1,0498	0,55287 - 1,5468
78-81	73	12	1,8056	1,19501 - 2,4161
81-85	62	27	0,8313	0,37937 - 1,2832
<hr/>				
19-57	172	12	2,6626	2,0774 - 3,2478
57-74	325	64	1,6249	1,3569 - 1,8930
74-81	187	50	1,3191	1,0070 - 1,6311
81-85	62	27	0,8313	0,3794 - 1,2832
<hr/>				
19-57	172	12	2,6626	2,0774 - 3,2478
57-85	574	141	1,4039	1,2196 - 1,5881

TAULA 4. Taula dels logaritmes dels odds de morir per a diverses categoritzacions de l'edat.

La figura 1 presenta els logaritmes dels odds per a les tres categoritzacions més la línia que millor aproxima els logaritmes dels odds per als decils. La línia que aproxima els dècils és equivalent al supòsit de linealitat dels OR . Sembla una simplificació coherent. També podríem utilitzar la categorització amb dues categories.

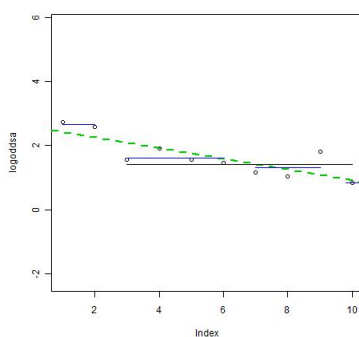


FIGURA 1. Els logaritmes dels odds segons diverses categoritzacions. La línia verda representa l'aproximació mínim quadrat als logaritmes dels odds segons els decils.

El model complet amb l'edat categoritzada en dues categories és

	Coef	S.E.	Wald Z	P
Intercept	-0.52667	5.69478	-0.09	0.9263
TEMPB	-0.09950	0.15524	-0.64	0.5216
TREAT=Citicoline	-0.09017	0.20027	-0.45	0.6525
SEX=Male	0.00530	0.20156	0.03	0.9790
RACE=B	0.38804	0.26699	1.45	0.1461
RACE=0	-0.97718	0.53747	-1.82	0.0690
BNIH	0.17872	0.02033	8.79	0.0000
decAB2=1	-1.10690	0.33839	-3.27	0.0011

L'edat en dues categories és significativa. Si contraposem els diversos estadístics de validació del model anterior i el definitiu tenim

Estadístic de validació	Model edat continu		Model edat en 2 categories	
	Valor	<i>p</i> -valor	Valor	<i>p</i> -valor
Raó de versemblança	116,970	$\leq 0,0001$	112.47	$\leq 0,0001$
D de Sommers	0,563		0.546	
Gamma	0,565		0.548	
Tau de Kendall	0,157		0.152	
R^2	0,216		0.209	
Area de la corba ROC	0,780		0,773	
Test de Hosmer-L		<i>p</i> -valor		<i>p</i> -valor
y=1		0,0033		0,0209

TAULA 5. Estadístics de validació i capacitat predictiva dels dos models logístic per a la mort del pacient.

El model amb l'edat millora el test de Hosmer-Lemeshow, però no el suficient com per a què compensi la introducció d'un nou paràmetre dintre del model.

Apèndix C

Taules de la linealitat dels cumulative logits de tenir seqüeles

Les taules següents presenten a cada fila la distribució de l'escala *mRS* per a cada octil de les variables numèriques. Els diversos cumulative *OR* quantifiquen la relació entre l'octil *i* i l'anterior, *i* - 1.

Octil	0-2	3-4	5	6	OR (3-6,0-2)	IC(OR,95%)	OR (5-6,0-4)	IC(OR,95%)	OR (6,0-5)	IC(OR,95%)
34,60-36,10	41	41	4	24						
36,10-36,30	41	42	4	13	0,855	[0,491; 1,490]	0,600	[0,305; 1,179]	0,535	[0,256; 1,120]
36,30-36,50	38	43	6	14	1,152	[0,654; 2,030]	1,206	[0,590; 2,465]	1,077	[0,479; 2,424]
36,50-36,72	47	41	7	16	0,821	[0,473; 1,425]	1,059	[0,541; 2,071]	1,047	[0,483; 2,270]
36,72-36,90	37	27	5	18	0,992	[0,563; 1,751]	1,375	[0,710; 2,665]	1,549	[0,738; 3,251]
36,90-37,11	43	36	8	18	1,067	[0,600; 1,898]	0,916	[0,478; 1,756]	0,793	[0,389; 1,639]
37,11-37,42	46	33	9	15	0,859	[0,496; 1,489]	0,923	[0,489; 1,744]	0,824	[0,391; 1,738]
37,42-39,80	25	42	2	18	2,001	[1,093; 3,667]	0,983	[0,499; 1,933]	1,530	[0,720; 3,254]

TAULA 1. Taula dels cumulative *OR* de tenir seqüeles al canviar el decil de la temperatura basal.

Octil	0-2	3-4	5	6	OR (3-6,0-2)	IC(OR,95%)	OR (5-6,0-4)	IC(OR,95%)	OR (6,0-5)	IC(OR,95%)
7-9	124	45	2	4						
9-10	37	22	1	4	1,774	[0,980; 3,212]	2,387	[0,702; 8,112]	2,850	[0,691; 11,753]
10-11	48	26	1	9	1,028	[0,532; 1,984]	1,595	[0,517; 4,920]	1,800	[0,528; 6,132]
11-13	50	40	2	13	1,467	[0,823; 2,613]	1,233	[0,523; 2,907]	1,178	[0,477; 2,905]
13-16	52	81	7	31	2,080	[1,258; 3,440]	1,714	[0,891; 3,300]	1,567	[0,779; 3,153]
16-17	10	34	7	25	2,884	[1,375; 6,049]	2,546	[1,424; 4,550]	2,214	[1,195; 4,102]
17-20	7	41	9	21	1,537	[0,553; 4,272]	0,859	[0,451; 1,637]	0,752	[0,376; 1,502]
20-34	5	37	18	39	1,854	[0,565; 6,082]	2,171	[1,185; 3,980]	1,764	[0,928; 3,354]

TAULA 2. Taula dels cumulative *OR* de tenir seqüeles al canviar el decil del NIHSS basal.

Octil	0-2	3-4	5	6	OR (3-6,0-2)	IC(OR,95%)	OR (5-6,0-4)	IC(OR,95%)	OR (6,0-5)	IC(OR,95%)
19,0-52,0	66	42	0	6						
52,0-60,0	48	44	1	15	1,719	[1,010; 2,924]	3,130	[1,177; 8,329]	2,903	[1,083; 7,786]
60,0-66,0	56	43	2	13	0,829	[0,489; 1,405]	0,871	[0,408; 1,862]	0,798	[0,361; 1,766]
66,0-70,0	33	35	7	17	1,726	[0,984; 3,030]	2,329	[1,139; 4,763]	1,761	[0,806; 3,847]
70,0-74,0	47	40	10	20	0,833	[0,474; 1,464]	0,977	[0,524; 1,823]	0,910	[0,446; 1,856]
74,0-77,0	27	40	9	23	1,791	[1,006; 3,186]	1,385	[0,767; 2,502]	1,468	[0,751; 2,870]
77,0-80,5	32	52	13	25	1,055	[0,580; 1,919]	0,947	[0,536; 1,674]	0,852	[0,449; 1,617]
80,5-85,0	24	30	5	27	0,919	[0,494; 1,708]	1,310	[0,732; 2,343]	1,776	[0,943; 3,344]

TAULA 3. Taula dels cumulative *OR* de tenir seqüeles al canviar el decil de l'edat.

Apèndix D

Codi R

En aquesta última secció s'inclou tot el codi R utilitzat.

1. Descriptiva General

```
source('bbdd.r')
#Taula basal
#funció auxiliar
desc<-function(var)
{
print('Mean Sd Median IQR')
cat(mean(var,na.rm=T),' ',sd(var,na.rm=T),' ',median(var,na.rm=T),
' ',IQR(var,na.rm=T),' \n')
}
with(subset(bbdd,bbdd$VISIT==0),desc(AGE))
with(subset(bbdd,bbdd$VISIT==0),desc(BNIH))
with(subset(bbdd,bbdd$VISIT==0),describe(TREAT))
with(subset(bbdd,bbdd$VISIT==0),describe(SEX))
with(subset(bbdd,bbdd$VISIT==0),describe(RACE))

#Figura histograma de la edad basal
with(bbdd,describe(AGE))
jpeg(filename = "histedat.jpg")
hist(bbdd[bbdd$VISIT==0,'AGE'],xlab='Edat (anys)',ylab='Freqüència', main='Edat')
abline(v=quantile(bbdd[bbdd$VISIT==0,'AGE'],c(0.025,0.5,0.975),na.rm=T),
col=c(2,3,2),lwd=2,lty=2)
dev.off()

#Figura histograma del NIH basal
jpeg(filename = "histnih.jpg")
hist(bbdd[bbdd$VISIT==0,'BNIH'],xlab='NIHSS basal',ylab='Freqüència',
main='NIHSS basal')
abline(v=quantile(bbdd[bbdd$VISIT==0,'BNIH'],c(0.025,0.5,0.975),na.rm=T),
```

```
col=c(2,3,2),lwd=2,lty=2)
dev.off()
```

2. Hipotèrmia induïda pel tractament

2.1. Descriptiva.

```
library(lattice)
source('bbdd.r')
options(digits=5)

#Histograma de totes les temperatres
jpeg('histtemp.jpeg',width=800,height=400)
par(mfrow=c(1,2))
hist(avit[, 'TEMPC'],xlab='Temp (°C)',main='Temperatura',xlim=c(34,40),
ylab='Frequència')
abline(v=quantile(avit[, 'TEMPC'],c(0.025,0.5,0.975),na.rm=T),col=c(2,3,2),
lwd=2,lty=2)
with(bbdd,qqnorm(TEMPC))
with(bbdd,qqline(TEMPC,col=2,lwd=2,lty=2))
dev.off()

#Evolució temporal
jpeg('evoltemp.jpeg')
boxplot(I(TEMPC-TEMPB)~VISIT,bbdd,subset=bbdd$VISIT>0,xlab='Visita',
ylab='diferència resp basal (°C)')
dev.off()
#with(bbdd,tapply(TEMPC,VISIT,desc1))
with(subset(bbdd,bbdd$VISIT>0),tapply(I(TEMPC-TEMPB),VISIT,desc1))

#Evolució de la temperatura per pacient
jpeg('evolpat.jpeg',width=1000,height=600)
xyplot(I(TEMPC-TEMPB)~VISIT|PATID,bbdd,subset=bbdd$VISIT>0,layout=c(12,4,1),
type="o",xlab='VISITA',ylab='Diferència resp basal (°C)')
dev.off()
jpeg('evolsegui.jpeg',width=1000,height=600)
xyplot(I(TEMPC-TEMPB)~VS.DAY|PATID,bbdd,subset=bbdd$VISIT>0,layout=c(12,4,1),
type="o",xlab='VISITA',ylab='Diferència resp basal (°C)',xlim=c(0,16*7))
dev.off()

#Evolució temporal segons el tractament
jpeg('evoltrac.jpeg',width=960,height=480)
boxplot(I(TEMPC-TEMPB)~TREAT+VISIT,bbdd,subset=bbdd$VISIT>0,col=c(2,3),
ylab='Diferència resp basal(°C)',axes=F)
axis(2)
axis(1,at=1:12,labels=c(paste(c('Pla','Citi'),0),paste(c('Pla','Citi'),1),
paste(c('Pla','Citi'),2),paste(c('Pla','Citi'),3),
paste(c('Pla','Citi'),6),paste(c('Pla','Citi'),12)))
```

```

box()
dev.off()
by((bbdd[, 'TEMPC']-bbdd[, 'TEMPB']),bbdd[,c('TREAT', 'VISIT')],desc1)

#Evolució temporal segons el sex
jpeg('evolsex.jpeg',width=960,height=480)
boxplot(I(TEMPC-TEMPB)~SEX+VISIT,bbdd,subset=bbdd$VISIT>0,col=c(2,3),
ylab='Diferència resp basal (°C)')
dev.off()
by(bbdd[, 'TEMPC']-bbdd[, 'TEMPB'],bbdd[,c('SEX', 'VISIT')],desc1)

#Evolució temporal segons la raça
jpeg('evolraca.jpeg',width=960,height=480)
boxplot(I(TEMPC-TEMPB)~RACE+VISIT,bbdd,subset=bbdd$VISIT>0,col=2:4,
ylab='Diferència resp basal (°C)',axes=F)
axis(2)
axis(1,at=1:(3*5),labels=c(paste(c('B', 'N', 'A'),1),
paste(c('B', 'N', 'A'),2),paste(c('B', 'N', 'A'),3),paste(c('B', 'N', 'A'),6),
paste(c('B', 'N', 'A'),12)))
box()
abline(v=3.5+3*0:5,lty=2)
dev.off()
latex(by(bbdd[, 'TEMPC']-bbdd[, 'TEMPB'],bbdd[,c('RACE', 'VISIT')],desc1))

#Evolució temporal segons visita i edat
jpeg('tempedat.jpeg',width=960,height=480)
xyplot(I(TEMPC-TEMPB)~AGE|VISIT,bbdd,subset=bbdd$VISIT>0,layout=c(5,1),
xlab='Edat (anys)',ylab='Diferència resp basal (°C)')
dev.off()

#Evolució temporal del BNIH segons visita
jpeg('tempnih.jpeg',width=960,height=480)
xyplot(I(TEMPC-TEMPB)~BNIH|VISIT,bbdd,subset=bbdd$VISIT>0,layout=c(5,1),
xlab='NIHSS',ylab='Diferència resp basal (°C)')
dev.off()

2.2. Model.
library(nlme)
source('bbdd.r')
long<-groupedData(I(TEMPC-TEMPB)~VS.DAY|PATID,bbdd)

#Model sense tractament
moddqua.def<-lme(I(TEMPC-TEMPB)~VS.DAY+I(VS.DAY^2)+SEX:VS.DAY,long,
subset=long$VISIT>0,random=~VS.DAY+I(VS.DAY^2)|PATID,na.action=na.exclude,
weights=varConstPower(form=~VS.DAY),control=lmeControl(tolerance=1e-4,
msVerbose=T,maxIter=200,msMaxIter=200,msTol=1e-4))
summary(moddqua.def)
anova(moddqua.def)

```

```

moddtract<-update(moddqua.def, .~.+TREAT+TREAT:VS.DAY+TREAT:I(VS.DAY^2))
summary(moddtract)
anova(moddtract)

jpeg('aproxpac.jpeg',width=800,height=600)
par(mfrow=c(3,6),bty='o')
for(i in 19:(2*18))
{
plot(I(TEMPC-TEMPB)~VS.DAY,bbdd[bbdd$PATID==row.names(coef(moddqua.var12))[i],],
type='o',main=row.names(coef(moddqua.var12))[i],xlim=c(0,100),ylim=c(-3,0),
xlab='Seguiment',ylab='Dif resp basal')
lines(coef(moddqua.def)[i,1]+(coef(moddqua.def)[i,2]+
coef(moddqua.def)[i,4]*(bbdd$SEX[bbdd$PATID==
row.names(coef(moddqua.def))[i]][1]=='Male'))*(1:(16*7))+
coef(moddqua.def)[i,3]*(1:(16*7))^2,col=2+(bbdd$SEX[bbdd$PATID==
row.names(coef(moddqua.def))[i]][1]=='Male'))
}
dev.off()

```

3. La mort en funció de la temperatura basal

3.1. Descriptiva.

```

library(lattice)
source('bbdd.r')
#taula basal
with(subset(bbdd,bbdd$VISIT==0),table(DIED,TREAT))
chisq.test(with(subset(bbdd,bbdd$VISIT==0),table(DIED,TREAT)))
oddsratio(with(subset(bbdd,bbdd$VISIT==0),table(DIED,TREAT)))
with(subset(bbdd,bbdd$VISIT==0),table(DIED,SEX))
chisq.test(with(subset(bbdd,bbdd$VISIT==0),table(DIED,SEX)))
oddsratio(with(subset(bbdd,bbdd$VISIT==0),table(DIED,SEX)))
with(subset(bbdd,bbdd$VISIT==0),table(DIED,RACE))
chisq.test(with(subset(bbdd,bbdd$VISIT==0),table(DIED,RACE)))
oddsratio(with(subset(bbdd,bbdd$VISIT==0),table(DIED,RACE)))

#taula basal per a les variables enteres
with(subset(bbdd,bbdd$VISIT==0),tapply(TEMPB,DIED,desc1))
with(subset(bbdd,bbdd$VISIT==0),tapply(BNIH,DIED,desc1))
with(subset(bbdd,bbdd$VISIT==0),tapply(AGE,DIED,desc1))

jpeg('basalmortent.jpeg',width=800,height=500)
par(mfrow=c(1,3))
with(subset(bbdd,bbdd$VISIT==0),boxplot(TEMPB~DIED,col=2:3,main='Basal
temperature',ylab='TEMP (°C)'))
with(subset(bbdd,bbdd$VISIT==0),boxplot(BNIH~DIED,col=2:3,main='Basal
NIH',

```

```
ylab='NIH'))
with(subset(bbdd,bbdd$VISIT==0),boxplot(AGE~DIED,col=2:3,main='AGE',
ylab='Age (years)'))
dev.off()
```

```
#Comparació de les mitjanes en els dos grups en la visita basal
t.test(TEMPC~DIED,bbdd,subset=bbdd$VISIT==0)
t.test(BNIH~DIED,bbdd,subset=bbdd$VISIT==0)
t.test(AGE~DIED,bbdd,subset=bbdd$VISIT==0)
```

3.2. Model.

```
library(lattice)
library(MASS)
library(Design)
library(Epi)
source('bbdd.r')
options(digits=5)

#Linealitat de les contínues
#Temp Basal
with(subset(bbdd,bbdd$VISIT==0),quantile(TEMPB,probs=seq(0,1,0.111111),
na.rm=T))
bbdd$decTB[bbdd$VISIT==0]<-with(subset(bbdd,bbdd$VISIT==0),cut2(TEMPB,g=10))
(tdec<-with(subset(bbdd,bbdd$VISIT==0),table(decTB,DIED)))
tdec2<-cbind(tdec[1:9,1],tdec[2:10,1],tdec[1:9,2],tdec[2:10,2])
(ORtdec<-tdec2[,4]*tdec2[,1]/(tdec2[,3]*tdec2[,2]))
(ORtdec2<-tdec[2:10,2]*tdec[1,1]/(tdec[2:10,1]*tdec[1,2]))
ORitdec<-tdec2[,4]*tdec2[,1]/(tdec2[,3]*tdec2[,2])*exp(-1.96*sqrt(
1/tdec2[,4]+1/tdec2[,1]+1/tdec2[,3]+1/tdec2[,2]))
ORstdec<-tdec2[,4]*tdec2[,1]/(tdec2[,3]*tdec2[,2])*exp(+1.96*sqrt(
1/tdec2[,4]+1/tdec2[,1]+1/tdec2[,3]+1/tdec2[,2]))
tdec
data.frame(tdec2,ORtdec,ORitdec,ORstdec)

#Basal NIH
with(subset(bbdd,bbdd$VISIT==0),quantile(BNIH,probs=seq(0,1,0.111111),
na.rm=T))
bbdd$decNB[bbdd$VISIT==0]<-with(subset(bbdd,bbdd$VISIT==0),cut2(BNIH,g=10))
(ndec<-with(subset(bbdd,bbdd$VISIT==0),table(decNB,DIED)))
(ndec2<-cbind(ndec[1:9,1],ndec[2:10,1],ndec[1:9,2],ndec[2:10,2]))
(ORndec<-ndec2[,4]*ndec2[,1]/(ndec2[,3]*ndec2[,2]))
(ORndec2<-ndec[2:10,2]*ndec[1,1]/(ndec[2:10,1]*ndec[1,2]))
(ORindec<-ndec2[,4]*ndec2[,1]/(ndec2[,3]*ndec2[,2])*exp(-1.96*sqrt(
1/ndec2[,4]+1/ndec2[,1]+1/ndec2[,3]+1/ndec2[,2]))
(ORsndec<-ndec2[,4]*ndec2[,1]/(ndec2[,3]*ndec2[,2])*exp(+1.96*sqrt(
1/ndec2[,4]+1/ndec2[,1]+1/ndec2[,3]+1/ndec2[,2]))
ndec
data.frame(ndec2,ORndec,ORindec,ORSndec)
```

```

#Basal Edat
with(subset(bbdd,bbdd$VISIT==0),quantile(AGE,probs=seq(0,1,0.111111),
na.rm=T))
nA=10
bbdd$decAB[bbdd$VISIT==0]<-with(subset(bbdd,bbdd$VISIT==0),cut2(AGE,g=nA))
(adec<-with(subset(bbdd,bbdd$VISIT==0),table(decAB,DIED)))
(adec2<-cbind(adec[1:(nA-1),1],adec[2:nA,1],adec[1:(nA-1),2],adec[2:nA,2]))
(ORadec<-adec2[,4]*adec2[,1]/(adec2[,3]*adec2[,2]))
(ORadec2<-adec[2:nA,2]*adec[1,1]/(adec[2:nA,1]*adec[1,2]))
(ORiadec<-adec2[,4]*adec2[,1]/(adec2[,3]*adec2[,2])*exp(-1.96*sqrt(
1/adec2[,4]+1/adec2[,1]+1/adec2[,3]+1/adec2[,2])))
(ORsadec<-adec2[,4]*adec2[,1]/(adec2[,3]*adec2[,2])*exp(+1.96*sqrt(
1/adec2[,4]+1/adec2[,1]+1/adec2[,3]+1/adec2[,2])))
adec
data.frame(adec2,ORadec,ORiadec,ORsadec)

jpeg('lineal.jpeg',width=800,height=400)
par(mfrow=c(1,3))
errbar(c('d2:d1','d3:d2','d4:d3','d5:d4','d6:d5','d7:d6','d8:d7',
'd9:d8','d10:d9'),log(ORtdec),log(ORitdec),log(ORstdec),ylab='log(OR)',
xlab='',ylim=c(-1.5,2))
title(main='Temperatura basal')
abline(v=coef(modOT)[2]+c(0,-1,1)*1.96*sqrt(modOT$var[2,2]),col=2,
lty=c(1,2,2))
text(log(ORtdec),1:9,round(ORtdec,2),pos=3)
ORndec[1]<-0
ORindec[1]<-0
ORsndec[1]<-0
errbar(c('d2:d1','d3:d2','d4:d3','d5:d4','d6:d5','d7:d6','d8:d7',
'd9:d8','d10:d9'),log(ORndec),log(ORindec),log(ORSndec),
ylab='log(OR)',xlab='',ylim=c(-1.5,2))
title(main='NIHSS basal')
abline(v=coef(modONB)[2]+c(0,-1,1)*1.96*sqrt(modONB$var[2,2]),col=2,
lty=c(1,2,2))
text(log(ORndec[2:9]),2:9,round(ORndec[2:9],2),pos=3)
errbar(c('d2:d1','d3:d2','d4:d3','d5:d4','d6:d5','d7:d6','d8:d7',
'd9:d8','d10:d9'),log(ORadec),log(ORiadec),log(ORSadec),ylab='log(OR)',xlab='',
ylim=c(-1.5,2))
title(main='Edat')
abline(v=coef(modOA)[2]+c(0,-1,1)*1.96*sqrt(modOA$var[2,2]),col=2,
lty=c(1,2,2))
text(log(ORadec),1:9,round(ORadec,2),pos=3)
dev.off()

#OR cru modOT<-lrm(DIED~TREAT,bbdd,subset=bbdd$VISIT==0)
exp(coef(modOT)[2]+c(0,-1,1)*1.96*sqrt(modOT$var[2,2]))
modOS<-lrm(DIED~SEX,bbdd,subset=bbdd$VISIT==0)
exp(coef(modOS)[2]+c(-1,0,1)*1.96*sqrt(modOS$var[2,2]))

```

```

modOR<-lrm(DIED~as.factor(RACE),bbdd,subset=bbdd$VISIT==0)
exp(coef(modOR)[2]+c(-1,0,1)*1.96*sqrt(modOR$var[2,2]))
exp(coef(modOR)[3]+c(-1,0,1)*1.96*sqrt(modOR$var[3,3]))
modOTB<-lrm(DIED~TEMPB,bbdd,subset=bbdd$VISIT==0)
exp(coef(modOTB)[2]+c(-1,0,1)*1.96*sqrt(modOTB$var[2,2]))
modONB<-lrm(DIED~BNIH,bbdd,subset=bbdd$VISIT==0)
exp(coef(modONB)[2]+c(-1,0,1)*1.96*sqrt(modONB$var[2,2]))
modOA<-lrm(DIED~AGE,bbdd,subset=bbdd$VISIT==0)
exp(coef(modOA)[2]+c(-1,0,1)*1.96*sqrt(modOA$var[2,2]))

#Model
moddef<-lrm(DIED~TEMPB+TREAT+SEX+as.factor(RACE)+BNIH+AGE,bbdd,
subset=bbdd$VISIT==0)
moddef
#OR ajustat
#TEMPC
exp(coef(moddef)[2]+c(0,-1,1)*1.96*sqrt(moddef$var[2,2]))
#TREAT
exp(coef(moddef)[3]+c(0,-1,1)*1.96*sqrt(moddef$var[3,3]))
#SEX
exp(coef(moddef)[4]+c(0,-1,1)*1.96*sqrt(moddef$var[4,4]))
#RACE
exp(coef(moddef)[5]+c(-1,0,1)*1.96*sqrt(moddef$var[5,5]))
exp(coef(moddef)[6]+c(0,-1,1)*1.96*sqrt(moddef$var[6,6]))
#BNIH
exp(coef(moddef)[7]+c(0,-1,1)*1.96*sqrt(moddef$var[7,7]))
#AGE
exp(coef(moddef)[8]+c(0,-1,1)*1.96*sqrt(moddef$var[8,8]))

#Mesures de la bondat
amoddef<-lrm(DIED~TEMPB+TREAT+SEX+as.factor(RACE)+BNIH+AGE,bbdd,
subset=bbdd$VISIT==0,x=T,y=T)
resid(amoddef,'gof')
#Curva ROC
jpeg('rocmort.jpeg')
aux<-subset(bbdd,bbdd$VISIT==0 & is.na(bbdd$TEMPB)==F)
with(aux,ROC(form=aux$DIED~TEMPB+TREAT+SEX+RACE+BNIH+AGE,plot='ROC',
xlab='1-especificitat'))
dev.off()
anova(moddef)
#var sig: BNIH i AGE
summary(moddef)

```

4. La temperatura basal com a factor pronòstic

4.1. Descriptiva.

```

library(lattice)
source('bbdd.r')

# Taula basal per a la nova categorització de mRS
with(subset(bbdd,bbdd$VISIT==12),table(RANKIN2))
with(subset(bbdd,bbdd$VISIT==12),table(TREAT,RANKIN2))
apply(with(subset(bbdd,bbdd$VISIT==12),table(TREAT,RANKIN2)),1,sum)
with(subset(bbdd,bbdd$VISIT==12),table(TREAT,RANKIN2)/423*100)
with(subset(bbdd,bbdd$VISIT==12),table(TREAT,RANKIN2)/429*100)
chisq.test(with(subset(bbdd,bbdd$VISIT==12),table(TREAT,RANKIN2)))
with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2))
apply(with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2)),1,sum)
with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2)/411*100)
with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2)/441*100)
chisq.test(with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2)))
apply(with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2)),1,sum)
round(with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2))/411*100,1)
round(with(subset(bbdd,bbdd$VISIT==12),table(SEX,RANKIN2))/441*100,1)
with(subset(bbdd,bbdd$VISIT==12),table(RACE,RANKIN2))
chisq.test(with(subset(bbdd,bbdd$VISIT==12),table(RACE,RANKIN2)))
apply(with(subset(bbdd,bbdd$VISIT==12),table(RACE,RANKIN2)),1,sum)
round(with(subset(bbdd,bbdd$VISIT==12),table(RACE,RANKIN2))/678*100,1)
round(with(subset(bbdd,bbdd$VISIT==12),table(RACE,RANKIN2))/127*100,1)
round(with(subset(bbdd,bbdd$VISIT==12),table(RACE,RANKIN2))/47*100,1)

# Taula basal per a les numèriques
with(subset(bbdd,bbdd$VISIT==12),tapply(TEMPB,RANKIN2,desc1))
with(subset(bbdd,bbdd$VISIT==12),tapply(BNIH,RANKIN2,describe))
with(subset(bbdd,bbdd$VISIT==12),tapply(AGE,RANKIN2,describe))

#Gràfics
jpeg('basalseq.jpeg',width=800,height=500)
par(mfrow=c(1,3))
with(subset(bbdd,bbdd$VISIT==12),boxplot(TEMPB~RANKIN2,col=2:8,main='Basal
temperature',ylab='TEMP (°C)',xlab='mRS'))
with(subset(bbdd,bbdd$VISIT==12),boxplot(BNIH~RANKIN2,col=2:8,main='Basal
NIH',
ylab='NIH',xlab='mRS'))
with(subset(bbdd,bbdd$VISIT==12),boxplot(AGE~RANKIN2,col=2:8,main='AGE',
ylab='Edat (anys)',xlab='mRS'))
dev.off()

```

4.2. Model.

```

library(lattice)
library(MASS)

```



```

library(Design)
library(Epi)
source('bbdd.r')
options(digits=5)

#supòsit de cumulative
#Temp Basal
with(subset(bbdd,bbdd$VISIT==12),quantile(TEMPB,probs=seq(0,1,0.125),na.rm=T))
bbdd$decTB[bbdd$VISIT==12]<-with(subset(bbdd,bbdd$VISIT==12),cut2(TEMPB,g=8))
(tdec<-with(subset(bbdd,bbdd$VISIT==12),table(decTB,RANKIN2)))
ORtdec<-matrix(0,nrow=7,ncol=3)
ORitdec<-ORtdec
ORstdec<-ORtdec
ORtdec[,1]<-apply(tdec[2:8,2:4],1,sum)*tdec[1:7,1]/
(tdec[2:8,1]*apply(tdec[1:7,2:4],1,sum))
ORitdec[,1]<-apply(tdec[2:8,2:4],1,sum)*tdec[1:7,1]/
(tdec[2:8,1]*apply(tdec[1:7,2:4],1,sum))*exp(-1.96*sqrt(
1/apply(tdec[2:8,2:4],1,sum)+1/tdec[1:7,1]+1/tdec[2:8,1]+
1/apply(tdec[1:7,2:4],1,sum)))
ORstdec[,1]<-apply(tdec[2:8,2:4],1,sum)*tdec[1:7,1]/
(tdec[2:8,1]*apply(tdec[1:7,2:4],1,sum))*exp(+1.96*sqrt(
1/apply(tdec[2:8,2:4],1,sum)+1/tdec[1:7,1]+1/tdec[2:8,1]+
1/apply(tdec[1:7,2:4],1,sum)))
ORtdec[,2]<-apply(tdec[2:8,3:4],1,sum)*apply(tdec[1:7,1:2],1,sum)/
(apply(tdec[2:8,1:2],1,sum)*apply(tdec[1:7,3:4],1,sum))
ORitdec[,2]<-apply(tdec[2:8,3:4],1,sum)*apply(tdec[1:7,1:2],1,sum)/
(apply(tdec[2:8,1:2],1,sum)*apply(tdec[1:7,3:4],1,sum))*
exp(-1.96*sqrt(1/apply(tdec[2:8,3:4],1,sum)+1/apply(tdec[1:7,1:2],1,sum)+
1/apply(tdec[2:8,1:2],1,sum)+1/apply(tdec[1:7,3:4],1,sum)))
ORstdec[,2]<-apply(tdec[2:8,3:4],1,sum)*apply(tdec[1:7,1:2],1,sum)/
(apply(tdec[2:8,1:2],1,sum)*apply(tdec[1:7,3:4],1,sum))*
exp(+1.96*sqrt(1/apply(tdec[2:8,3:4],1,sum)+1/apply(tdec[1:7,1:2],1,sum)+
1/apply(tdec[2:8,1:2],1,sum)+1/apply(tdec[1:7,3:4],1,sum)))
ORtdec[,3]<-tdec[2:8,4]*apply(tdec[1:7,1:3],1,sum)/
(apply(tdec[2:8,1:3],1,sum)*tdec[1:7,4])
ORitdec[,3]<-tdec[2:8,4]*apply(tdec[1:7,1:3],1,sum)/
(apply(tdec[2:8,1:3],1,sum)*tdec[1:7,4])*
exp(-1.96*sqrt(1/tdec[2:8,4]+1/apply(tdec[1:7,1:3],1,sum)+
1/apply(tdec[2:8,1:3],1,sum)+1/tdec[1:7,4]))
ORstdec[,3]<-tdec[2:8,4]*apply(tdec[1:7,1:3],1,sum)/
(apply(tdec[2:8,1:3],1,sum)*tdec[1:7,4])*
exp(+1.96*sqrt(1/tdec[2:8,4]+1/apply(tdec[1:7,1:3],1,sum)+
1/apply(tdec[2:8,1:3],1,sum)+1/tdec[1:7,4]))
tdec
data.frame(ORtdec,ORitdec,ORstdec)
#Basal NIH
with(subset(bbdd,bbdd$VISIT==12),quantile(BNIH,probs=seq(0,1,0.125),na.rm=T))
bbdd$decNB[bbdd$VISIT==12]<-with(subset(bbdd,bbdd$VISIT==12),cut2(BNIH,g=8))

```

```

(ndec<-with(subset(bbdd,bbdd$VISIT==12),table(decNB,RANKIN2)))
ORndec<-matrix(0,nrow=7,ncol=3)
ORindec<-ORndec
ORSndec<-ORndec
(ORndec[,1]<-apply(ndec[2:8,2:4],1,sum)*ndec[1:7,1]/
(ndec[2:8,1]*apply(ndec[1:7,2:4],1,sum)))
ORindec[,1]<-apply(ndec[2:8,2:4],1,sum)*ndec[1:7,1]/
(ndec[2:8,1]*apply(ndec[1:7,2:4],1,sum))*exp(-1.96*sqrt(
1/apply(ndec[2:8,2:4],1,sum)+1/ndec[1:7,1]+1/ndec[2:8,1]+
1/apply(ndec[1:7,2:4],1,sum)))
ORSndec[,1]<-apply(ndec[2:8,2:4],1,sum)*ndec[1:7,1]/
(ndec[2:8,1]*apply(ndec[1:7,2:4],1,sum))*exp(+1.96*sqrt(
1/apply(ndec[2:8,2:4],1,sum)+1/ndec[1:7,1]+1/ndec[2:8,1]+
1/apply(ndec[1:7,2:4],1,sum)))
ORndec[,2]<-apply(ndec[2:8,3:4],1,sum)*apply(ndec[1:7,1:2],1,sum)/
(apply(ndec[2:8,1:2],1,sum)*apply(ndec[1:7,3:4],1,sum))
ORindec[,2]<-apply(ndec[2:8,3:4],1,sum)*apply(ndec[1:7,1:2],1,sum)/
(apply(ndec[2:8,1:2],1,sum)*apply(ndec[1:7,3:4],1,sum))*
exp(-1.96*sqrt(1/apply(ndec[2:8,3:4],1,sum)+1/apply(ndec[1:7,1:2],1,sum)+
1/apply(ndec[2:8,1:2],1,sum)+1/apply(ndec[1:7,3:4],1,sum)))
ORSndec[,2]<-apply(ndec[2:8,3:4],1,sum)*apply(ndec[1:7,1:2],1,sum)/
(apply(ndec[2:8,1:2],1,sum)*apply(ndec[1:7,3:4],1,sum))*
exp(+1.96*sqrt(1/apply(ndec[2:8,3:4],1,sum)+1/apply(ndec[1:7,1:2],1,sum)+
1/apply(ndec[2:8,1:2],1,sum)+1/apply(ndec[1:7,3:4],1,sum)))
ORndec[,3]<-ndec[2:8,4]*apply(ndec[1:7,1:3],1,sum)/
(apply(ndec[2:8,1:3],1,sum)*ndec[1:7,4])
ORindec[,3]<-ndec[2:8,4]*apply(ndec[1:7,1:3],1,sum)/
(apply(ndec[2:8,1:3],1,sum)*ndec[1:7,4])*
exp(-1.96*sqrt(1/ndec[2:8,4]+1/apply(ndec[1:7,1:3],1,sum)+
1/apply(ndec[2:8,1:3],1,sum)+1/ndec[1:7,4]))
ORSndec[,3]<-ndec[2:8,4]*apply(ndec[1:7,1:3],1,sum)/
(apply(ndec[2:8,1:3],1,sum)*ndec[1:7,4])*
exp(+1.96*sqrt(1/ndec[2:8,4]+1/apply(ndec[1:7,1:3],1,sum)+
1/apply(ndec[2:8,1:3],1,sum)+1/ndec[1:7,4]))
ndec[1,]
data.frame(ndec[2:8,],cbind(ORndec,ORindec,ORSndec)[,c(1,4,7,2,5,8,3,6,9)])
#Basal Edat
with(subset(bbdd,bbdd$VISIT==12),quantile(AGE,probs=seq(0,1,0.125),na.rm=T))
bbdd$decAB[bbdd$VISIT==12]<-with(subset(bbdd,bbdd$VISIT==12),cut2(AGE,g=8))
(adec<-with(subset(bbdd,bbdd$VISIT==12),table(decAB,RANKIN2)))
ORadec<-matrix(0,nrow=7,ncol=3)
ORiadec<-ORadec
ORSadec<-ORadec
(ORadec[,1]<-apply(adec[2:8,2:4],1,sum)*adec[1:7,1]/
(adec[2:8,1]*apply(adec[1:7,2:4],1,sum)))
ORiadec[,1]<-apply(adec[2:8,2:4],1,sum)*adec[1:7,1]/
(adec[2:8,1]*apply(adec[1:7,2:4],1,sum))*exp(-1.96*sqrt(
1/apply(adec[2:8,2:4],1,sum)+1/adec[1:7,1]+1/adec[2:8,1]+

```

```

1/apply(adec[1:7,2:4],1,sum)))
ORsadec[,1]<-apply(adec[2:8,2:4],1,sum)*adec[1:7,1]/
(adec[2:8,1]*apply(adec[1:7,2:4],1,sum))*exp(+1.96*sqrt(
1/apply(adec[2:8,2:4],1,sum)+1/adec[1:7,1]+1/adec[2:8,1]+
1/apply(adec[1:7,2:4],1,sum)))
ORadec[,2]<-apply(adec[2:8,3:4],1,sum)*apply(adec[1:7,1:2],1,sum)/
(apply(adec[2:8,1:2],1,sum)*apply(adec[1:7,3:4],1,sum))
ORiadec[,2]<-apply(adec[2:8,3:4],1,sum)*apply(adec[1:7,1:2],1,sum)/
(apply(adec[2:8,1:2],1,sum)*apply(adec[1:7,3:4],1,sum))*
exp(-1.96*sqrt(1/apply(adec[2:8,3:4],1,sum)+1/apply(adec[1:7,1:2],1,sum)+
1/apply(adec[2:8,1:2],1,sum)+1/apply(adec[1:7,3:4],1,sum)))
ORsadec[,2]<-apply(adec[2:8,3:4],1,sum)*apply(adec[1:7,1:2],1,sum)/
(apply(adec[2:8,1:2],1,sum)*apply(adec[1:7,3:4],1,sum))*
exp(+1.96*sqrt(1/apply(adec[2:8,3:4],1,sum)+1/apply(adec[1:7,1:2],1,sum)+
1/apply(adec[2:8,1:2],1,sum)+1/apply(adec[1:7,3:4],1,sum)))
ORadec[,3]<-adec[2:8,4]*apply(adec[1:7,1:3],1,sum)/
(apply(adec[2:8,1:3],1,sum)*adec[1:7,4])
ORiadec[,3]<-adec[2:8,4]*apply(adec[1:7,1:3],1,sum)/
(apply(adec[2:8,1:3],1,sum)*adec[1:7,4])*
exp(-1.96*sqrt(1/adec[2:8,4]+1/apply(adec[1:7,1:3],1,sum)+
1/apply(adec[2:8,1:3],1,sum)+1/adec[1:7,4]))
ORsadec[,3]<-adec[2:8,4]*apply(adec[1:7,1:3],1,sum)/
(apply(adec[2:8,1:3],1,sum)*adec[1:7,4])*
exp(+1.96*sqrt(1/adec[2:8,4]+1/apply(adec[1:7,1:3],1,sum)+
1/apply(adec[2:8,1:3],1,sum)+1/adec[1:7,4]))
adec
data.frame(ORadec,ORiadec,ORsadec)
adec[1,]
data.frame(adec[2:8,],cbind(ORadec,ORiadec,ORsadec)[,c(1,4,7,2,5,8,3,6,9)])

library(gplots)
jpeg('mrslineal.jpeg',width=1000,height=600)
noms=c('d2:d1','d3:d2','d4:d3','d5:d4','d6:d5','d7:d6','d8:d7')
cex.lab=(1)
period=rep(seq(0.5,6.5,1),3)
par(mfrow=c(1,3))
OR=log(c(ORtdec))
lower=log(c(ORitdec))
l=OR-lower
upper=log(c(ORstdec))
u=upper-OR
plotCI(period,OR,uiv=u,liw=l,xlab="",ylab='log(OR)',add=F,ylim=c(-1.5,2.5),
col=rep(1:3,each=7),barcol=rep(1:3,each=7),lwd=2,sfrac=0.03,axes=F,
main='Temperatura basal',xlim=c(0,7))
axis(2)
axis(1,at=period[1:7],labels=noms)
legend('topright',legend=c('m6-3:m2-0','m6-5:m4-0','m6:m5-0'),col=1:3,lty=1)
abline(h=coef(mod0tb)[4]+c(-1,0,1)*1.96*sqrt(mod0tb$var[4,4]),col=4,

```

```

lty=c(2,1,2))
text(period,OR,round(c(ORtdec),2),pos=c(rep(2,7),
rep(3,3),1,2,3,3,2,1,2,2,1,4,2),col=rep(1:3,each=7))
OR=log(c(ORndec))
lower=log(c(ORindec))
l=OR-lower
upper=log(c(ORsndec))
u=upper-OR
plotCI(period,OR,uiv=u,liw=1,xlab="",ylab='log(OR)',add=F,ylim=c(-1.5,2.5),
col=rep(1:3,each=7),barcol=rep(1:3,each=7),lwd=2,sfrac=0.03,axes=F,
main='NIHSS basal',xlim=c(0,7))
axis(2)
axis(1,at=period[1:7],labels=noms)
legend('topright',legend=c('m6-3:m2-0','m6-5:m4-0','m6:m5-0'),col=1:3,lty=1)
abline(h=coef(mod0bn)[4]+c(-1,0,1)*1.96*sqrt(mod0bn$var[4,4]),col=4,
lty=c(2,1,2))
text(period,OR,round(c(ORtdec),2),pos=c(rep(2,6),4,
rep(2,7),2,2,4,4,2,2,2),col=rep(1:3,each=7))
OR=log(c(ORadec))
lower=log(c(ORiadec))
l=OR-lower
upper=log(c(ORsadec))
u=upper-OR
plotCI(period,OR,uiv=u,liw=1,xlab="",ylab='log(OR)',add=F,ylim=c(-1.5,2.5),
col=rep(1:3,each=7),barcol=rep(1:3,each=7),lwd=2,sfrac=0.03,axes=F,main='Edat',
xlim=c(0,7))
axis(2)
axis(1,at=period[1:7],labels=noms)
legend('topright',legend=c('m6-3:m2-0','m6-5:m4-0','m6:m5-0'),col=1:3,lty=1)
abline(h=coef(mod0ed)[4]+c(-1,0,1)*1.96*sqrt(mod0ed$var[4,4]),col=4,
lty=c(2,1,2))
text(period,OR,round(c(ORtdec),2),pos=c(2,4,rep(2,4),4,
2,3,2,3,2,4,2,4,2,4,4,4,2,2),col=rep(1:3,each=7),csi=12)
dev.off()

#Canvi d'octil pel cumulative
jpeg('mrslineal2.jpeg',width=10*300,height=4.5*300,res=300)
noms=c('d2:d1','d3:d2','d4:d3','d5:d4','d6:d5','d7:d6','d8:d7')
par(mfrow=c(1,3))
#TEMPB
nTB=8
bbdd$decTB[bbdd$VISIT==12]<-with(subset(bbdd,bbdd$VISIT==12),cut2(TEMPB,g=nTB))
mod0tempb<-lrm(as.factor(RANKIN2)~as.factor(decTB),
data=bbdd,subset=bbdd$VISIT==12,x=T,y=T)
mod0tempb
OR=coef(mod0tempb)[4:10]
l=1.96*sqrt(diag(mod0tempb$var)[4:10])
u=1.96*sqrt(diag(mod0tempb$var)[4:10])

```

```

plotCI(4:10,OR,uIW=u,liw=1,xlab="",ylab='log(OR)',add=F,ylim=c(-1,4),
col=1,barcol=1,lwd=2,sfrac=0.03,axes=F,
main='Temperatura basal',xlim=c(3.5,10.5))
abline(h=coef(mod0tb)[4]+c(-1,0,1)*1.96*sqrt(mod0tb$var[4,4]),col=4,
lty=c(2,1,2))
axis(2)
axis(1,at=4:10,labels=noms)
text(4:10,OR,round(exp(OR),2),pos=4,col=1)
box()

#BNIH
nNB=8
quantile(bbdd$BNIH[bbdd$VISIT==12],probs=seq(0,1,1/nNB),na.rm=T)
bbdd$decNB[bbdd$VISIT==12]<-with(subset(bbdd,bbdd$VISIT==12),cut2(BNIH,g=nNB))
table(bbdd$decNB[bbdd$VISIT==12])
mod0nihb<-lrm(as.factor(RANKIN2)~as.factor(decNB),
data=bbdd,subset=bbdd$VISIT==12,x=T,y=T)
mod0nihb
(OR=coef(mod0nihb)[4:(3+(nNB-1))])
l=1.96*sqrt(diag(mod0nihb$var)[4:(3+(nNB-1))])
u=1.96*sqrt(diag(mod0nihb$var)[4:(3+(nNB-1))])
plotCI(4:(3+(nNB-1)),OR,uIW=u,liw=1,xlab="",ylab='log(OR)',add=F,ylim=c(-1,4),
col=1,barcol=1,lwd=2,sfrac=0.03,axes=F,
main='NIHSS basal',xlim=c(3.5,10.5))
abline(h=coef(mod0tb)[4]+c(-1,0,1)*1.96*sqrt(mod0tb$var[4,4]),col=4,
lty=c(2,1,2))
axis(2)
axis(1,at=4:(3+(nNB-1)),labels=noms)
text(4:(3+(nNB-1)),OR,round(exp(OR),2),pos=4,col=1)
box()

#AGE
nAB=8
quantile(bbdd$AGE[bbdd$VISIT==12],probs=seq(0,1,1/nAB),na.rm=T)
bbdd$decAB[bbdd$VISIT==12]<-with(subset(bbdd,bbdd$VISIT==12),cut2(AGE,g=nAB))
table(bbdd$decAB[bbdd$VISIT==12])
mod0ageb<-lrm(as.factor(RANKIN2)~as.factor(decAB),
data=bbdd,subset=bbdd$VISIT==12,x=T,y=T)
mod0ageb
(OR=coef(mod0ageb)[4:(3+(nAB-1))])
l=1.96*sqrt(diag(mod0ageb$var)[4:(3+(nAB-1))])
u=1.96*sqrt(diag(mod0ageb$var)[4:(3+(nAB-1))])
plotCI(4:(3+(nAB-1)),OR,uIW=u,liw=1,xlab="",ylab='log(OR)',add=F,ylim=c(-1,4),
col=rep(1:3,each=(nAB-1)),barcol=rep(1:3,each=(nAB-1)),lwd=2,sfrac=0.03,axes=F,
main='Edat basal',xlim=c(3.5,10.5))
abline(h=coef(mod0ed)[4]+c(-1,0,1)*1.96*sqrt(mod0tb$var[4,4]),col=4,
lty=c(2,1,2))
axis(2)

```

```

axis(1,at=4:(3+(nAB-1)),labels=noms)
text(4:(3+(nAB-1)),OR,round(exp(OR),2),pos=4,col=1)
box()
dev.off()

mod1decNB<-lrm(as.factor(RANKIN2)~as.factor(decNB),data=bbdd,subset=bbdd$VISIT==12,x=T,y=T)
mod1decNB
round(cbind(exp(coef(mod1decNB)[4:10]),exp(coef(mod1decNB)[4:10]-
1.96*sqrt(diag(mod1decNB$var))[4:10]),exp(coef(mod1decNB)[4:10]+
1.96*sqrt(diag(mod1decNB$var))[4:10])),4)
mod1decAB<-lrm(as.factor(RANKIN2)~as.factor(decAB),data=bbdd,subset=bbdd$VISIT==12,x=T,y=T)
mod1decAB
round(cbind(exp(coef(mod1decAB)[4:10]),exp(coef(mod1decAB)[4:10]-
1.96*sqrt(diag(mod1decAB$var))[4:10]),exp(coef(mod1decAB)[4:10]+
1.96*sqrt(diag(mod1decAB$var))[4:10])),4)
mod1tot3<-lrm(as.factor(RANKIN2)~TEMPB+TREAT+SEX+as.factor(RACE)+
as.factor(decNB)+as.factor(decAB),
data=bbdd,subset=bbdd$VISIT==12,x=T,y=T)
mod1tot3
round(cbind(exp(coef(mod1tot3)[4:22]),exp(coef(mod1tot3)[4:22]-
1.96*sqrt(diag(mod1tot3$var))[4:22]),exp(coef(mod1tot3)[4:22]+
1.96*sqrt(diag(mod1tot3$var))[4:22])),4)
round(resid(mod1tot3,'gof'),6)
#efectes no ajustats
mod0tb<-lrm(as.factor(RANKIN2)~TEMPB,data=bbdd,subset=bbdd$VISIT==12)
exp(coef(mod0tb)[4]+c(-1,0,1)*1.96*sqrt(mod0tb$var[4,4]))
mod0tr<-lrm(as.factor(RANKIN2)~TREAT,data=bbdd,subset=bbdd$VISIT==12)
exp(coef(mod0tr)[4]+c(-1,0,1)*1.96*sqrt(mod0tr$var[4,4]))
mod0se<-lrm(as.factor(RANKIN2)~SEX,data=bbdd,subset=bbdd$VISIT==12)
exp(coef(mod0se)[4]+c(-1,0,1)*1.96*sqrt(mod0se$var[4,4]))
mod0ra<-lrm(as.factor(RANKIN2)~as.factor(RACE),data=bbdd,subset=bbdd$VISIT==12)
exp(coef(mod0ra)[4]+c(-1,0,1)*1.96*sqrt(mod0ra$var[4,4]))
exp(coef(mod0ra)[5]+c(-1,0,1)*1.96*sqrt(mod0ra$var[5,5]))
mod0bn<-lrm(as.factor(RANKIN2)~decBN,data=bbdd,subset=bbdd$VISIT==12)
round(exp(coef(mod0bn)[4:10]+c(-1,0,1)*1.96*sqrt(diag(mod0bn$var)[4:10])),4)
mod0ed<-lrm(as.factor(RANKIN2)~decAB,data=bbdd,subset=bbdd$VISIT==12)
round(exp(coef(mod0ed)[4:10]+c(-1,0,1)*1.96*sqrt(diag(mod0ed$var)[4:10])),4)

#model
mod1tot<-lrm(as.factor(RANKIN2)~TEMPB+TREAT+SEX+as.factor(RACE)+BNIH+AGE,
data=bbdd,subset=bbdd$VISIT==12,x=T,y=T)
mod1tot
anova(mod1tot)

#OR ajustat exp(coef(mod1tot)[4]+c(-1,0,1)*1.96*sqrt(mod1tot$var[4,4]))
exp(coef(mod1tot)[5]+c(-1,0,1)*1.96*sqrt(mod1tot$var[5,5]))
exp(coef(mod1tot)[6]+c(-1,0,1)*1.96*sqrt(mod1tot$var[6,6]))
exp(coef(mod1tot)[7]+c(-1,0,1)*1.96*sqrt(mod1tot$var[7,7]))

```

```

exp(coef(mod1tot)[8]+c(-1,0,1)*1.96*sqrt(mod1tot$var[8,8]))
exp(coef(mod1tot)[9]+c(-1,0,1)*1.96*sqrt(mod1tot$var[9,9]))
exp(coef(mod1tot)[10]+c(-1,0,1)*1.96*sqrt(mod1tot$var[10,10]))

#Mesures de bondat
resid(mod1tot,'gof')

```

5. Annex: Construcció del model sense tractament

```

library(lattice)
library(nlme)
source('bbdd.r')

#Model lineal
modlin<-lm(TEMPC~VS.DAY,bbdd)
summary(modlin)
plot(modlin)
jpeg('anlineal.jpeg')
plot(fitted(modlin),residuals(modlin),xlab='Valors estimats',ylab='Residus')
dev.off()
plot(bbdd$TEMPC[!is.na(bbdd$TEMPC)])
#no tenim un bon model

# Model lineal amb efectes aleatoris
modlinRE<-lme(TEMPC~VS.DAY,bbdd,random=~VS.DAY|PATID,na.exclude)
summary(modlinRE)
jpeg('anlinRE.jpeg')
plot(modlinRE,xlab='Valors estimats',ylab='Residus estandaritzats')
dev.off()

#Provem el model quadràtic amb efectes aleatoris lineals
modquaRE<-lme(TEMPC~VS.DAY+I(VS.DAY^2),bbdd,random=~VS.DAY|PATID,
na.action=na.exclude,control=lmeControl(tolerance=1e-4,maxIter=200,
msMaxIter=200,msTol=1e-4))
summary(modquaRE)
anova(update(modlinRE,method='ML'),update(modquaRE,method='ML'))
#millor el quadràtic
jpeg('anquaRE.jpeg')
plot(modquaRE,xlab='Valors estimats',ylab='Residus estandaritzats')
dev.off()

#Model quadratic amb efectes aleatoris quadràtics
modquaRE2<-update(modquaRE,random=~VS.DAY+I(VS.DAY^2)|PATID)
#no convergeix
#provem el model quadràtic amb efectes aleatoris quadràtics amb varPower
long<-groupedData(TEMPC~VS.DAY+I(VS.DAY^2)|PATID,bbdd)
modquaRE2<-lme(TEMPC~VS.DAY+I(VS.DAY^2),long,random=~VS.DAY+I(VS.DAY^2)|PATID,

```

```

na.action=na.exclude,weights=varPower(form=~VS.DAY),
control=lmeControl(tolerance=1e-4,maxIter=200,msMaxIter=200,msTol=1e-4))
summary(modquaRE2)
jpeg('anquaRE.jpeg')
plot(modquaRE2,xlab='Valors estimats',ylab='Residus estandaritzats')
dev.off()
anova(update(modlinRE,method='ML'),update(modquaRE2,method='ML'))

#Transformem les dades
bbdd2<-bbdd[bbdd$VISIT>0,]
long2<-groupedData(I(TEMPC-TEMPB)~VS.DAY+I(VS.DAY^2)|PATID,bbdd2)
#Model lineal amb efectes aleatoris
moddlinRE<-lme(I(TEMPC-TEMPB)~VS.DAY,bbdd2,na.action=na.exclude,
random=~VS.DAY|PATID)
summary(moddlinRE)
anova(moddlinRE)
jpeg('andlinRE.jpeg')
plot(moddlinRE,xlab='Valors estimats',ylab='Residus estandaritzats')
dev.off()

#Models lineals amb diverses varFunc
moddlinRE2<-update(moddlinRE,weights=varPower(form=~VS.DAY))
moddlinRE3<-update(moddlinRE,weights=varConstPower(form=~VS.DAY))
moddlinRE4<-update(moddlinRE,weights=varExp(form=~VS.DAY))
moddlinRE5<-update(moddlinRE,weights=varFixed(form=~VS.DAY))
#No convergeix el model amb varFixed
anova(moddlinRE,moddlinRE2,moddlinRE3,moddlinRE4)#,moddlinRE5)
anova(moddlinRE2,moddlinRE,moddlinRE4)
#millor model amb varExp

#model Quadràtic amb efectes aleatoris quadràtics
moddquaRE<-lme(I(TEMPC-TEMPB)~VS.DAY+I(VS.DAY^2),bbdd2,na.action=na.exclude,
random=~VS.DAY+I(VS.DAY^2)|PATID,
control=lmeControl(tolerance=1e-6,maxIter=200,msMaxIter=200,msTol=1e-6))
summary(moddquaRE)
anova(moddquaRE)
anova(update(moddlinRE,method='ML'),update(moddquaRE,method='ML'))
#Millor el quadràtic
jpeg('andquaRE.jpeg')
plot(moddquaRE,xlab='Valors estimats',ylab='Residus estandaritzats')
dev.off()
#Diferents varFunc
moddquaRE2<-update(moddquaRE,weights=varPower(form=~VS.DAY),
control=lmeControl(tolerance=1e-4,maxIter=200,msMaxIter=200,msTol=1e-4))
#No convergeix moddquaRE2<-update(moddquaRE,weights=varConstPower(form=~VS.DAY))
moddquaRE2ML<-update(moddquaRE,method='ML')
moddquaRE4<-update(moddquaRE,weights=varExp(form=~VS.DAY))
moddquaRE5<-update(moddquaRE,weights=varFixed(form=~VS.DAY))

```



```

#No convergeix
anova(moddquaRE,moddquaRE2,moddquaRE4)
#Millor amb varConstPower

#Comparem lineal i quadràtic
anova(update(moddquaRE,method='ML'),update(moddlinRE4,method='ML'),
update(moddquaRE2,method='ML'))
#millor el quadràtic amb varConstPower

#Les var d'estudi només entren en els fixos
#SEX
moddqua.var11<-update(moddquaRE2,.~.+SEX+SEX:VS.DAY+SEX:I(VS.DAY^2))
summary(moddqua.var11)
anova(moddqua.var11)
moddqua.var12<-update(moddquaRE2,.~.+SEX+SEX:VS.DAY)
summary(moddqua.var12)
anova(moddqua.var12)
moddqua.var13<-update(moddquaRE2,.~.+SEX:VS.DAY)
summary(moddqua.var13)
anova(moddqua.var13)
moddqua.var13ML<-update(moddqua.var13,method='ML')
anova(update(moddqua.var11,method='ML'),update(moddqua.var12,method='ML'),
moddqua.var13ML,moddquaRE2ML)
#Millor model amb només la interacció amb el terme lineal
#RACE
moddqua.var21<-update(moddqua.var13,.~.+RACE+RACE:VS.DAY+RACE:I(VS.DAY^2))
summary(moddqua.var21)
anova(moddqua.var21)
moddqua.var22<-update(moddqua.var13,.~.+RACE+RACE:I(VS.DAY^2))
summary(moddqua.var22)
anova(moddqua.var22)
moddqua.var23<-update(moddqua.var13,.~.+RACE)
summary(moddqua.var23)
anova(moddqua.var23)
anova(update(moddqua.var21,method='ML'),update(moddqua.var22,method='ML'),
update(moddqua.var23,method='ML'),moddqua.var13ML)

#BNIH
moddqua.var31<-update(moddqua.var13,.~.+BNIH+BNIH:VS.DAY+BNIH:I(VS.DAY^2))
summary(moddqua.var31)
anova(moddqua.var31)
moddqua.var32<-update(moddqua.var13,.~.+BNIH:VS.DAY+BNIH:I(VS.DAY^2))
summary(moddqua.var32)
anova(moddqua.var32)
moddqua.var32ML<-update(moddqua.var32,method='ML')
moddqua.var33<-update(moddqua.var13,.~.+BNIH:I(VS.DAY^2))
summary(moddqua.var33)
anova(moddqua.var33)

```

```

anova(update(moddqua.var31,method='ML'),moddqua.var32ML,
update(moddqua.var33,method='ML'),moddqua.var13ML)
anova(moddqua.var13ML,moddqua.var32ML)
#millor amb només el sexe

#AGE
moddqua.var41<-update(moddqua.var13,.~.+AGE+AGE:VS.DAY+AGE:I(VS.DAY^2))
summary(moddqua.var41)
anova(moddqua.var41)
moddqua.var42<-update(moddqua.var13,.~.+AGE+AGE:VS.DAY)
summary(moddqua.var42)
anova(moddqua.var42)
moddqua.var43<-update(moddqua.var13,.~.+AGE)
summary(moddqua.var43)
anova(update(moddqua.var41,method='ML'),update(moddqua.var42,method='ML'),
update(moddqua.var43,method='ML'),moddqua.var13ML)
moddqua.var41<-update(moddqua.var32,.~.+AGE+AGE:VS.DAY+AGE:I(VS.DAY^2))
summary(moddqua.var41)
anova(moddqua.var41)
moddqua.var42<-update(moddqua.var32,.~.+AGE+AGE:VS.DAY)
summary(moddqua.var42)
anova(moddqua.var42)
moddqua.var43<-update(moddqua.var32,.~.+AGE)
summary(moddqua.var43)
anova(moddqua.var43)
anova(update(moddqua.var41,method='ML'),update(moddqua.var42,method='ML'),
update(moddqua.var43,method='ML'),moddqua.var32ML)
anova(moddqua.var13ML,moddquaRE2ML,moddqua.var32ML)

plot(moddquaRE2)

sum(abs(resid(moddquaRE2,type='p'))> 1.96,na.rm=T)/length(resid(moddquaRE2,type='p'))*100
sum(abs(resid(moddqua.var13,type='p'))> 1.96,na.rm=T)/length(resid(moddqua.var13,type='p'))*100

```

6. Annex: Model amb pacients amb el tractament complet

```

library(lattice)
library(nlme)
source('bbdd.r')

#fem el model sobre els pacients amb com a molt una visita perduda
#executar el codi de tempc.r
indcom<-complete.cases(diagrama[,c(1,5:9)])
bbdds<-subset(bbdd,bbdd$PATID %in% levels(bbdd$PATID)[indcom])

```

```
moddef<-lme(I(TEMPC-TEMPB)~VS.DAY+I(VS.DAY^2),bbdds,
subset=bbdds$VISIT>0,na.action=na.exclude,random= VS.DAY+I(VS.DAY^2)|PATID,
weights=varConstPower(form= VS.DAY),
control=lmeControl(tolerance=1e-6,maxIter=200,msMaxIter=200,msTol=1e-6))
summary(moddef)
plot(moddef)

moddef2<-update(moddef,.~.+TREAT+TREAT:VS.DAY+TREAT:I(VS.DAY^2))
summary(moddef2)
moddef3<-update(moddef,.~.+TREAT:VS.DAY+TREAT:I(VS.DAY^2))
summary(moddef3)
```


Bibliografia

1. S. A. Bernard, T. W. Gray, M. D. Buist, B. M. Jones, W. Silvester, G. Gutteridge, and K. Smith, *Treatment of comatose survivors of out-of-hospital cardiac arrest with induced hypothermia*, N. Engl. J. Med. **346** (2002), 557–563.
2. W. M. Clark, S. J. Warach, L. C. Pettigrew, and R. E. Gammans, *A randomized dose-response trial of citicoline in acute ischemic stroke patients*, Neurology **49** (1997), no. 3, 671–678.
3. W. M. Clark, L. R. Wechsler, L. A. Sabounjian, and Schwiderski UE, *Citicoline stroke study group. a phase iii randomized efficacy trial of 2000 mg citicoline in acute ischemic stroke patients*, Neurology **57** (2001), no. 9, 1595–1602.
4. W. M. Clark, B. J. Williams, K. A. Selzer, R. M. Zweifler, L. A. Sabounjian, and R. E. Gammans, *A randomized efficacy trial of citicoline in patients with acute ischemic stroke*, Stroke **30** (1999), no. 12, 2592–2597.
5. A. Davalos, J. Castillo, J. Alvarez-Sabin, J. J. Secades, J. Mercadal, S. Lopez, E. Cobo, S. Warach, D. Sherman, W. M. Clark, and R. Lozano, *Oral citicoline in acute ischemic stroke: an individual patient data pooling analysis of clinical trials*, Stroke **33** (2002), no. 12, 2850–2857.
6. H. M. den Hertog, H. B. van der Worp, H. M. van Gemert, A. Algra, L. J. Kappelle, J. van Gijn, P.J. Koudstaal, and D. W. Dippel, *An early rise in body temperature is related to unfavorable outcome after stroke: data from the pais study*, J. Neurol. (2010).
7. P. Diggle, K-Y Liang, and S. L. Zeger, *Analysis of longitudinal data*, Oxford University Press, 2002.
8. D. W. J. Dippel, E. J. van Breda, H. M. A. van Gemert, H. B. van der Worp, R. J. Meijer, L. J. Kappelle, P. J. Koudstaal, and on Behalf of the PAPAS Investigators, *Effect of paracetamol (acetaminophen) on body temperature in acute ischemic stroke. a double-blind, randomized phase ii clinical trial*, Stroke **32** (2001), 1607–1612.
9. P. Farreras and C. Rozman (eds.), *Medicina interna*, 15 ed., Elsevier, 2004.
10. O. A. Harris, Jr J. M. Colford, M. C. Good, and P. G. Matz, *The role of hypothermia in the management of severe brain injury. a meta-analysis*, Arch Neurol. **59** (2002), 1077–1083.
11. M. Holzer, *Mild therapeutic hypothermia to improve the neurologic outcome after cardiac arrest*, N. Engl. J. Med. **346** (2002), no. 8, 549–556.
12. D. W. Hosmer and S. Lemeshow, *Applied logistic regression*, John Wiley and Sons, 2000.
13. L. P. Kammersgaard, H. S. Jorgensen, J. A. Rungby, J. Reith, H. Nakayama, U. J. Weber, J. Houth, and T. T. Olsen, *Admission body temperature predicts long-term mortality after acute stroke. the copenhagen stroke study*, Stroke **33** (2002), 1759–1762.
14. N. M. Laird and J. H. Ware, *Random-effects models for longitudinal data*, Biometrics **38** (1982), 963–974.
15. D. W. Marion, L. E. Penrod, S. F. Kelsey, W. D. Obrist, P. M. Kochanek, A. M. Palmer, S. R. Wisniewskim, and S. T. Dekosky, *Treatment of traumatic brain injury with moderate hypothermia*, N. Engl. J. Med. **336** (1997), 540–546.
16. R. J. Meijer, H. Visser, P. J. Koudstaal, and W. J. Dippel, *Lowering body temperature in acute ischemic stroke without artificial ventilation and heavy sedation: A feasibility study*, Journal of Stroke and Cerebrovascular Diseases **10** (2001), no. 4, 157–160.

17. M. Millán, L. Grau, M. Castellanos, M. Rodríguez-Yáñez, J. F. Arenillas, F. Nombela, N. Pérez de la Ossa, L. López-Manzanares, J. Serena, J. Castillo, and A. Dávalos, *Body temperature and response to thrombolytic therapy in acute ischaemic stroke*, *European Journal of Neurology* **15** (2008), 1384–1389.
18. J. C. Pinheiro, *Topics in mixed-effects models*, Ph.D. thesis, University of Wisconsin, 1994.
19. J. C. Pinheiro and D. M. Bates, *Mixed-effects models in s and s-plus*, Springer, 2000.
20. J. Reith, H.S. Jørgensen, H. Nakamaya P.M. Pedersen and, L.L. Jeppesen, and H.O. Raaschou T.S. Olsen and, *Body temperature in acute stroke: relation to stroke severity, infarct size, mortality, and outcome*, *The Lancet* **347** (1996), 422–425.
21. J. D. Singer and J. B. Willet, *Applied longitudinal data analysis: Modelling change and event occurrence*, Oxford University Press, New York, 2003.
22. G. Verbeke and G. Molenberghs, *Linear mixed models for longitudinal data*, Springer-Verlag, 2000.