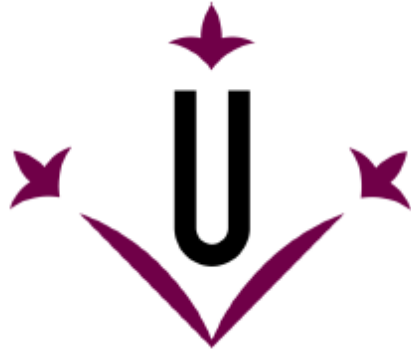


Universitat de Lleida



Anàlisis i valoració de resultats d'un estudi

Per: Jofre Illa Casals i Xavi Martinez Santamaria

Facultat de medicina

Doble Grau en Nutrició i Dietètica Humana i Fisioteràpia.

Treball presentat a: *Albert Sorribas Tello.*

Bioestadística.

Curs: 2019-2020.

Juny de 2020.

ÍNDEX:

INTRODUCCIÓ:.....	3
Anàlisi de les condicions inicials dels grups:	4
Model de relació univariant:	6
Model de relació multivariant: regressió logística:.....	8
Impacte de la diabetis:	11
La adherència a la dieta mediterrània:	14
ANÀLISI DE SUPERVIVÈNCIA	16
CONCLUSIONS:	21

INTRODUCCIÓ:

En el present document es recull una tasca d'avaluació relativa a la assignatura de Bioestadística inclosa en el doble grau de Nutrició i dietètica humana i Fisioteràpia que s'imparteix a la Universitat de Lleida.

En aquesta tasca s'encomana a l'alumnat el desenvolupament d'una anàlisi estadística i crítica sobre les dades referents a un estudi. En aquest cas concret, l'estudi en qüestió s'anomena PREDIMED: un estudi experimental que pretén obtenir una visió objectiva sobre la variable de la dieta humana (en concret la mediterrània amb complements com l'oli d'oliva verge extra o els fruits sec) com a element que desencadena un paper amb efecte significatiu sobre les malalties cardiovasculars o la mort derivada per aquesta tipologia de patologies.

Per tal de desenvolupar aquesta anàlisi s'han aplicat les diferents o més adequades tècniques o estratègies analítiques en estadística que s'han tractat i definit al llarg del trimestre universitari en el transcurs de la assignatura.

Anàlisi de les condicions inicials dels grups:

L'estudi PREDIMED (Prevenió amb dieta mediterrània) és del tipus multi-cèntric i es va desenvolupar a Espanya amb un total de 7447 participants que disposaven d'un risc elevat de patir una malaltia cardiovascular però que, tot i el pretext, no havia desenvolupat cap malaltia d'aquesta tipologia. En aquest sentit, es tracta de subjectes que presenten diabetis tipus 2 o tres dels factors de major risc com serien: fumar, hipertensió, LDL elevat, HDL reduït, sobrepès o obesitat o història prèvia familiar sobre malalties coronàries.

Aquests participants van ser assignats aleatòriament a una de les tres dietes que contemplava l'estudi com a condicions experimentals: dieta mediterrània amb suplementació d'oli d'oliva verge extra, dieta mediterrània amb fruits secs o una dieta control (amb una reducció de greixos de la dieta). A partir d'aquesta assignació randomitzada, els grups van ser seguits, de forma prospectiva, durant diversos anys (els autors determinen una mitja de seguiment dels participants d'entorn als 4,8 anys).

En aquest punt inicial, és interessant aturar-se a analitzar les condicions de partida o d'inici que disposaven els diferents grups un cop establerta aquesta aleatorització. Abans de profunditzar en la anàlisi de la taula de comparació de grups, cal puntualitzar que es tracta d'un estudi amb una mostra de població elevada (n=7447), la qual cosa permet una potència estadística elevada que permet detectar, de forma significativa, les diferències mínimes que es puguin donar en la distribució de grups que va resultar de la aleatorització:

-----Summary descriptives table by 'Intervention group'-----

	Control N=2042	MedDiet + Nuts N=2100	MedDiet + VO0 N=2182	p.overall
AMI, stroke, or CV Death:				0.064
No	1945 (95.2%)	2030 (96.7%)	2097 (96.1%)	
Yes	97 (4.75%)	70 (3.33%)	85 (3.90%)	
Sex:				<0.001
Male	812 (39.8%)	968 (46.1%)	899 (41.2%)	
Female	1230 (60.2%)	1132 (53.9%)	1283 (58.8%)	
Age	67.3 (6.28)	66.7 (6.02)	67.0 (6.21)	0.003
Smoking:				0.444
Never	1282 (62.8%)	1259 (60.0%)	1351 (61.9%)	
Current	270 (13.2%)	296 (14.1%)	292 (13.4%)	
Former	490 (24.0%)	545 (26.0%)	539 (24.7%)	
Body mass index	30.3 (3.96)	29.7 (3.77)	29.9 (3.71)	<0.001
Waist circumference	101 (10.8)	100 (10.6)	100 (10.4)	0.045
Waist-to-height ratio	0.63 (0.07)	0.62 (0.06)	0.63 (0.06)	<0.001
Hypertension:				0.249
No	331 (16.2%)	362 (17.2%)	396 (18.1%)	
Yes	1711 (83.8%)	1738 (82.8%)	1786 (81.9%)	
Type-2 diabetes:				0.017
No	1072 (52.5%)	1150 (54.8%)	1100 (50.4%)	
Yes	970 (47.5%)	950 (45.2%)	1082 (49.6%)	
Dyslipidemia:				0.423
No	563 (27.6%)	561 (26.7%)	622 (28.5%)	
Yes	1479 (72.4%)	1539 (73.3%)	1560 (71.5%)	
Family history of premature CHD:				0.581
No	1580 (77.4%)	1640 (78.1%)	1675 (76.8%)	
Yes	462 (22.6%)	460 (21.9%)	507 (23.2%)	
Hormone-replacement therapy:				0.850
No	1811 (98.3%)	1835 (98.4%)	1918 (98.2%)	
Yes	31 (1.68%)	30 (1.61%)	36 (1.84%)	
MedDiet Adherence score	8.44 (1.94)	8.81 (1.90)	8.77 (1.97)	<0.001

De la taula anterior es pot determinar una descriptiva sobre les condicions inicials de cada grup segons les variables establertes per l'estudi:

- Sobre la població de mostra de cada grup: en termes generals, els tres grups presenten una mostra similar en quant a nombre d'individus que la componen.
- Esdeveniment: No s'observen diferències significatives per a aquesta variable en aquesta taula descriptiva.
 - Els grups presenten major població sana en comparativa a la població que desenvolupa l'esdeveniment.
 - La presència d'esdeveniment és lleugerament superior en el grup control.
- Sobre el sexe (gènere): El p-valor indica diferències significatives entre els grups que poden relacionar-se a:
 - Variacions en el nombre d'individus per a cada condició experimental respecte a les variables home o dona.
 - Una distribució major de dones en els tres grups respecte a la població d'homes.
- Edat: El p-valor indica diferències significatives entre els grups:
 - Les edats de mitjana d'entre els grups són força similars. No obstant això, la potència estadística per l'amplitud de mostra pot ocasionar que la diferència de mesos en la mitjana d'edat de cada grup suposi una dada significativa.
- Hàbit de fumar: Aquesta dada no presenta diferències significatives entre els grups pels diferents nivells de la variable. En aquest sentit, doncs, els tres grups presenten una distribució no diferencial en quant a nombre de persones que fumen, han deixat de fumar o mai han fumat.
- Índex de massa corporal: Aquesta dada presenta diferències significatives entre grups segons el p-valor de la dada. Valorant les mitjanes d'aquesta variable per grup es pot observar que, en els tres casos, aquesta es situa en valors de sobrepès o lleugera obesitat. En aquesta dada, tal com es descriu, no diferencia entre altres variables que puguin orientar la descriptiva (tipus de massa corporal [si és grassa o magra], estatura de la persona, gènere, etc.
- Circumferència de la cintura i rati entre cintura i alçada: El p-valor indica que hi ha diferències significatives entre els grups en quant a aquestes variables.
- Hipertensió: Aquesta variable no presenta diferències significatives entre els grups indicant una bona distribució entre si referent a població amb o sense hipertensió.
- Diabetis tipus 2: La dada es descriu amb diferències significatives entre els grups:
 - S'observa que, en el grup referent a la dieta mediterrània amb complement d'oli d'oliva verge extra, hi ha una població lleugerament major de persones amb diabetis.
- Dislipèmia: en la descriptiva s'observa una distribució simètrica dels individus per a aquesta variable.
- Historia familiar prèvia de malaltia cardiovascular prematura: en la descriptiva no s'observen diferències significatives en quant a la distribució de les variables entre els grups.
- Teràpia de reemplaçament hormonal: Per a aquesta variable no s'observen diferències significatives en els grups.
- Resultat d'adherència a la dieta mediterrània: Per a aquesta variable hi ha diferències significatives entre grups tot i que els resultats que es reflecteixen en les taules són força ajustats entre si.
- Mitjana d'anys de seguiment: Les diferències són significatives segons s'especifica a la taula:
 - S'observa un seguiment lleugerament major en el grup de dieta mediterrània amb suplement d'oli d'oliva verge extra.

Model de relació uni-variant:

Un cop feta la anàlisi de les condicions inicials i disposar d'una descriptiva que permet valorar una igualtat entre els grups per a les diferents variables incloses a l'inici de l'estudi com a factors que poden influir en el desenllaç que es pretén investigar, es valora adequadament iniciar una anàlisi uni-variant centrada en la observació de les odds ratio per a cada variable respecte al resultat objecte d'estudi:

-----Summary descriptives table by 'AMI, stroke, or CV Death'-----

	No N=6072	Yes N=252	OR	p.ratio	p.overall
Intervention group:					0.064
Control	1945 (32.0%)	97 (38.5%)	Ref.	Ref.	
MedDiet + Nuts	2030 (33.4%)	70 (27.8%)	0.69 [0.50;0.95]	0.021	
MedDiet + VOO	2097 (34.5%)	85 (33.7%)	0.81 [0.60;1.09]	0.173	
Sex:					<0.001
Male	2528 (41.6%)	151 (59.9%)	Ref.	Ref.	
Female	3544 (58.4%)	101 (40.1%)	0.48 [0.37;0.62]	<0.001	
Age	66.9 (6.14)	69.4 (6.65)	1.07 [1.04;1.09]	<0.001	<0.001
Smoking:					<0.001
Never	3778 (62.2%)	114 (45.2%)	Ref.	Ref.	
Current	809 (13.3%)	49 (19.4%)	2.01 [1.41;2.82]	<0.001	
Former	1485 (24.5%)	89 (35.3%)	1.99 [1.49;2.64]	<0.001	
Body mass index	30.0 (3.81)	29.8 (3.92)	0.98 [0.95;1.02]	0.365	0.378
Waist circumference	100 (10.6)	102 (10.6)	1.01 [1.00;1.03]	0.016	0.017
Waist-to-height ratio	0.63 (0.07)	0.63 (0.07)	3.64 [0.55;23.9]	0.178	0.187
Hypertension:					0.879
No	1047 (17.2%)	42 (16.7%)	Ref.	Ref.	
Yes	5025 (82.8%)	210 (83.3%)	1.04 [0.75;1.48]	0.826	
Type-2 diabetes:					<0.001
No	3231 (53.2%)	91 (36.1%)	Ref.	Ref.	
Yes	2841 (46.8%)	161 (63.9%)	2.01 [1.55;2.62]	<0.001	
Dyslipidemia:					<0.001
No	1645 (27.1%)	101 (40.1%)	Ref.	Ref.	
Yes	4427 (72.9%)	151 (59.9%)	0.56 [0.43;0.72]	<0.001	
Family history of premature CHD:					0.403
No	4694 (77.3%)	201 (79.8%)	Ref.	Ref.	
Yes	1378 (22.7%)	51 (20.2%)	0.87 [0.63;1.18]	0.363	
Hormone-replacement therapy:					0.186
No	5341 (98.2%)	223 (99.6%)	Ref.	Ref.	
Yes	96 (1.77%)	1 (0.45%)	0.29 [0.01;1.27]	0.117	
MeDiet Adherence score	8.70 (1.94)	8.24 (1.94)	0.89 [0.84;0.95]	<0.001	<0.001
follow-up to main event (years)	4.41 (1.66)	2.96 (1.76)	0.60 [0.55;0.65]	<0.001	<0.001

De la taula anterior se'n poden extreure diverses observacions:

- Respecte al tipus d'intervenció: com a variable no presenta, a priori, un valor de significança en la implicació respecte a episodis de malaltia cardiovascular o mort per aquest tipus de malaltia.
 - En el grup que segueix una dieta mediterrània amb suplementació de fruits secs s'observa una odds de 0.69 amb un IC (0.50, 0.95) que permetria valorar-la com a protectora davant aquestes patologies.
 - Pel que fa a la dieta mediterrània suplementada amb oli d'oliva verge extra, presenta una odds de 0.81 amb un IC (0.60, 1.09) que no permet descartar la igualtat de resultats respecte al grup control.
- Sexe (gènere): pel que fa a aquesta dada, s'observa un risc major de patir aquestes malalties en els homes respecte a les dones, segons la mostra observada.
- Edat: sembla que l'augment de l'edat pot comportar, a la llarga, un augment del risc de patir aquestes malalties.
- Hàbit de fumar: Els resultats de les odds ratio per a aquesta variable són significatives:

- El fet de disposar de l'hàbit de fumar augmenta, de mitja, un 2.01 el risc de patir aquesta malalties amb un IC (1.41, 2.82).
- Per a les persones que han deixat de fumar aquest risc és lleugerament menor: 1.99 amb IC (1.49, 2.64).
- Índex de massa corporal: segons aquesta anàlisi estadística, aquesta variable no es relaciona amb un risc significatiu per a patir aquesta tipologia de malalties: OR: 0.98 amb IC (0.95, 1.02).
- Circumferència de maluc: Aquesta variable no presenta un valor significatiu en quant a l'augment o disminució del risc de patir aquesta tipologia de malalties.
- Rati circumferència de maluc-alçada: Aquesta variable no presenta un valor significatiu en quant a l'augment o disminució del risc de patir aquesta tipologia de malalties. Alhora, es tracta d'una variable que presenta una gran incertesa: IC: 0.55, 23.9.
- Hipertensió: La taula aporta un valor no significatiu en quant al risc que pot originar-se d'aquesta variable. L'OR s'estableix en 1.04 amb un IC: 0.75, 1.48. Per tant, no es pot concloure que les persones amb hipertensió presentin major risc de patir malalties cardiovasculars respecte als qui no disposen d'aquest factor.
- Diabetis tipus 2: Aquest factor suposa clarament una variable de risc pel desenvolupament de malalties o mort per causes cardiovasculars per a aquelles persones que són diabètiques.
- Dislipèmia: Aquesta variable es postula com a factor que pot jugar cert paper protector davant les malalties càrdio-vasculars.
- Història familiar de malaltia coronària prematura: els resultats d'aquesta variable per aquesta anàlisi indiquen que, el fet disposar d'aquesta història prèvia, no suposa un risc significatiu en el desenllaç de malalties cardiovasculars o mort per dites causes.
- Teràpia hormonal de reemplaçament: Aquesta variable no presenta un resultat d'OR significatiu.
- Adherència a la dieta mediterrània: aquesta pot suposar un factor protector (OR: 0.89 amb IC: 0.84, 0.95) davant el risc de patir malalties cardiovasculars.

Model de relació multivariant: regressió logística:

Un cop feta la anàlisi uni-variant centrada en l'observació de les odds ratio, s'han utilitzat el mètode lineal generalitzat per tal de poder descartar aquelles variables que no siguin significatives pel que fa a patir malaltia cardiovascular.

En primer lloc, tal com es pot veure en la imatge següent, s'ha aplicat el model entre "event", que en el nostre cas es refereix a patir una malaltia cardiovascular, i les dues dietes que es proporcionen, deixant a banda totes les altres variables.

En aquest cas veiem que la dieta basada en fruits secs té un p-valor molt baix i per tant que s'ha d'incloure en aquest model.

```
Call:
glm(formula = event ~ group, family = "binomial", data = predimed)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.3120 -0.3120 -0.2819 -0.2604  2.6081

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)    -2.9983    0.1040  -28.820  <2e-16 ***
groupMedDiet + Nuts  -0.3690    0.1600   -2.306   0.0211 *
groupMedDiet + VOO  -0.2073    0.1519   -1.365   0.1723
---
                        2.5 % 97.5 %
(Intercept)           0.05 0.04  0.06
groupMedDiet + Nuts  0.69 0.50  0.94
groupMedDiet + VOO  0.81 0.60  1.09
```

A més a més, també hem aplicat un model amb totes les variables que hem exposat anteriorment, afegint també els dos tipus de dietes.

Es pot observar com diferents variables tenen un p-valor elevat:

- En primer lloc s'ha de destacar el p-valor elevat (0,17) de la variable que fa referència a la dieta rica en oli d'oliva verge extra. Anteriorment ja havíem vist com la dieta rica en fruits secs era més determinant però en aquest cas, en comparació amb totes les altres variables, podem veure com n'hi ha d'altres que són més significatives que no pas la dieta rica en VOO.
- La variable "bmi", que es refereix a l'índex de massa corporal, veiem que té un p-valor= 0,25 conclouent que aquesta variable no tindrà un paper important en l'aparició de malalties cardiovasculars.
- Veiem com la variable "famhistYes" també té un p-valor que s'aproxima molt a 1. Per tant els antecedents familiars no jugaran un paper important.
- La variable de teràpia hormonal de reemplaçament també es descarta del model ja que el seu p-valor no és significatiu.


```

Call:
glm(formula = event ~ group + sex + age + smoke + bmi + waist +
     wth + htn + diab + hyperchol + famhist + hormo + p14, family = "binomial",
     data = predimed)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7874 -0.3124 -0.2360 -0.1779  3.0769

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -7.30040    1.19599  -6.104 1.03e-09 ***
groupMedDiet + Nuts -0.40748    0.17479  -2.331 0.019740 *
groupMedDiet + VOO -0.22389    0.16306  -1.373 0.169743
sexFemale     -0.80325    0.23271  -3.452 0.000557 ***
age           0.06506    0.01193   5.456 4.88e-08 ***
smokeCurrent  0.57703    0.23148   2.493 0.012673 *
smokeFormer   0.43774    0.19755   2.216 0.026699 *
bmi          -0.03417    0.02970  -1.150 0.249960
waist        -0.03088    0.01846  -1.673 0.094393 .
wth          7.32502    3.03404   2.414 0.015767 *
htnYes       0.44260    0.20377   2.172 0.029853 *
diabYes      0.55788    0.15301   3.646 0.000266 ***
hypercholYes -0.33897    0.14691  -2.307 0.021039 *
famhistYes   0.18858    0.17406   1.083 0.278614
hormoYes     -0.96434    1.01590  -0.949 0.342496
p14          -0.10547    0.03535  -2.983 0.002850 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Per tant, s'hauria d'aplicar un model nou descartant les variables mencionades anteriorment ja que en aquest cas no eren significatives. Un cop descartades aquestes cinc variables es torna a veure que en aquest nou model, anomenat "Model 1", apareixen dues variables amb p-valors massa elevats:

- En primer lloc destacar la variable que fa referència a la mesura de la cintura, que amb un p-valor de 0,66 queda totalment descartada.
- En aquest model també es pot observar que el fet de patir hipertensió no serà significatiu en aquest cas.

MODEL 1

```

Call:
glm(formula = event ~ group + sex + age + smoke + waist + htn +
     diab + hyperchol + p14, family = "binomial", data = predimed)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8742 -0.3134 -0.2413 -0.1857  3.0253

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -7.422468    1.098052  -6.760 1.38e-11 ***
groupMedDiet + Nuts -0.309994    0.162969  -1.902 0.057149 .
groupMedDiet + VOO -0.166220    0.154446  -1.076 0.281822
sexFemale     -0.454732    0.176326  -2.579 0.009911 **
age           0.068701    0.010719   6.409 1.46e-10 ***
smokeCurrent  0.700388    0.211533   3.311 0.000930 ***
smokeFormer   0.482028    0.184602   2.611 0.009023 **
waist         0.002821    0.006534   0.432 0.665993
htnYes       0.285406    0.181312   1.574 0.115461
diabYes      0.559364    0.143855   3.888 0.000101 ***
hypercholYes -0.303184    0.138973  -2.182 0.029139 *
p14          -0.112408    0.032822  -3.425 0.000615 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

```

Com ja s'havia fet anteriorment, com que el Model 1 contenia variables no significatives, es necessita calcular un nou model per poder veure si totes les variables restants són significatives. En aquest cas, aquest model tindrà el nom de "Model 2", on hauran quedat descartades tant la variable cintura com la d'hipertensió.

I finalment, tal com es pot veure a la taula adjuntada, la dieta rica en fruits secs, el sexe femení, l'edat, ser fumador actual, ser antic fumador, patir diabetis, tenir hipercolesterolèmia i p14 seran variables significatives en patir malalties cardiovasculars.

MODEL 2

```
Call:
glm(formula = event ~ group + sex + age + smoke + diab + hyperchol +
     p14, family = "binomial", data = predimed)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.8699  -0.3138  -0.2405  -0.1864   3.0227

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -6.92734    0.83074  -8.339 < 2e-16 ***
groupMedDiet + Nuts -0.31916    0.16278  -1.961 0.049920 *
groupMedDiet + VOO -0.17815    0.15422  -1.155 0.248013
sexFemale    -0.45493    0.17378  -2.618 0.008847 **
age           0.06999    0.01068   6.556 5.52e-11 ***
smokeCurrent  0.67723    0.21040   3.219 0.001288 **
smokeFormer   0.47966    0.18373   2.611 0.009039 **
diabYes       0.51095    0.14110   3.621 0.000293 ***
hypercholYes -0.29323    0.13906  -2.109 0.034965 *
p14           -0.11587    0.03263  -3.551 0.000384 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Seguidament s'ha realitzat una taula d'OR de les variables seleccionades al Model 2.

```
> cbind(coef(model.2)%>%exp()%>%round(2),
+        confint(model.2)%>%exp()%>%round(2))
Waiting for profiling to be done...
              2.5 % 97.5 %
(Intercept)  0.00  0.00  0.00
groupMedDiet + Nuts 0.73 0.53 1.00
groupMedDiet + VOO 0.84 0.62 1.13
sexFemale     0.63 0.45 0.89
age           1.07 1.05 1.10
smokeCurrent  1.97 1.30 2.96
smokeFormer   1.62 1.13 2.32
diabYes       1.67 1.27 2.20
hypercholYes  0.75 0.57 0.98
p14           0.89 0.84 0.95
```

- En primer lloc trobem la dieta rica en fruits secs, que si ens fixem amb el valor de l'OR (0,73), podem arribar a la conclusió que ajudarà a prevenir malalties cardiovasculars.

- La variable del sexe femení té un OR de 0,63. Això significa que el sexe femení presenta menor

freqüència en aquesta tipologia de patologies. Dit d'una altra manera, que els homes seran més propensos a patir aquest tipus de malalties.

- L'edat té un OR molt semblant a 1 i per tant, en aquest cas, no podem dir que l'augment de l'edat estigui relacionat amb l'augment de probabilitats de patir la malaltia.
- Pel que fa a fumadors habituals, tenen un OR de 1'97. És a dir, el fet de ser fumador habitual augmenta les possibilitats de patir malalties cardiovasculars.
- En el cas de la variable de pacients que havien fumat anteriorment, té una relació directament proporcional amb el fet de patir la malaltia, tot i que amb un OR inferior al de la variable de fumadors habituals.
- El fet de patir diabetis també augmenta el risc, com en el cas anterior.
- En canvi, en aquest model el fet de patir hipercolesterolèmia té un OR de 0'75, tenint un efecte protector.
- L'adherència a la dieta mediterrània té un OR de 0'89, amb un valor inferior a 1 que fa que sigui una variable protectora.

Impacte de la diabetis:

A continuació es fa una anàlisi de la variable diabetis com a factor que presenta, en el cas de patir dita malaltia degenerativa, una implicació important en el desenvolupament de l'event relacionat a malalties càrdio-vasculars com les que analitza l'estudi Predimed. Així doncs, l'interès d'aquest punt es centra en determinar quines variables incloses en l'estudi Predimed estan directament relacionades amb la presència de diabetis en els participants inclosos en l'estudi per tal de detectar variables o casuístiques que permetin determinar factors que puguin associar-se a una patologia com la diabetis i, com a factor de risc, a un posterior desenllaç en malaltia càrdio-vascular.

D'aquesta manera, doncs, es desenvolupa l'anàlisi d'un model per a determinar les variables relacionades amb la diabetis:

- En aquest primer model:

```
Call:
glm(formula = diab ~ sex + age + smoke + bmi + waist + htn +
     hyperchol + famhist, family = "binomial", data = predimed)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3383 -1.0009 -0.5975  1.0471  2.2966

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.229180  0.435732   0.526 0.598912
sexFemale   -0.195619  0.078268  -2.499 0.012442 *
age          0.017477  0.004661   3.749 0.000177 ***
smokeCurrent -0.945304  0.099309  -9.519 < 2e-16 ***
smokeFormer -0.203213  0.081553  -2.492 0.012710 *
bmi         -0.043202  0.010932  -3.952 7.75e-05 ***
waist        0.025770  0.003999   6.443 1.17e-10 ***
htnYes      -1.562421  0.083744 -18.657 < 2e-16 ***
hypercholYes -1.344619  0.065566 -20.508 < 2e-16 ***
famhistYes  -0.952267  0.070564 -13.495 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8750.7 on 6323 degrees of freedom
Residual deviance: 7545.6 on 6314 degrees of freedom
AIC: 7565.6

Number of Fisher Scoring iterations: 4
```

- Es relaciona la variable d'interès diabetis amb variables que poden estar associades a la presència, ja existent a l'inici de l'estudi, de diabetis en els participants.
- Es destaquen com les variables sexe, estat de consum de tabac (consum actual o passat), índex de massa corporal, hipertensió, hipercolesterolèmia, historia familiar prèvia de malaltia càrdio-vascular.
- Es planteja un nou model amb aquestes variables. S'exclouran les variables maluc (waist) i teràpia hormonal (hormoYes).

- En el segon model:

```
Call:
glm(formula = diab ~ sex + age + smoke + bmi + htn + hyperchol +
     famhist, family = "binomial", data = predimed)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4005  -1.0082  -0.6183   1.0476   2.2438

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.205753   0.407455   2.959 0.00308 **
sexFemale    -0.370135   0.073328  -5.048 4.47e-07 ***
age           0.019669   0.004632   4.247 2.17e-05 ***
smokeCurrent -0.923751   0.099012  -9.330 < 2e-16 ***
smokeFormer  -0.178596   0.081195  -2.200 0.02784 *
bmi           0.008321   0.007438   1.119 0.26325
htnYes       -1.552487   0.083299 -18.638 < 2e-16 ***
hypercholYes -1.341930   0.065321 -20.544 < 2e-16 ***
famhistYes   -0.949921   0.070315 -13.510 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8750.7 on 6323 degrees of freedom
Residual deviance: 7588.0 on 6315 degrees of freedom
AIC: 7606

Number of Fisher Scoring iterations: 4
```

○ En aquest darrer model es pot observar que la variable relativa a l'índex de massa corporal (bmi) esdevé no significativa per a la anàlisi efectuat.

○ La resta de variables es mantenen amb associació respecte a la diabetis.

- El tercer model:

```
Call:
glm(formula = diab ~ sex + age + smoke + htn + hyperchol + famhist,
     family = "binomial", data = predimed)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4067  -1.0096  -0.6253   1.0462   2.2471

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.474654   0.329237   4.479 7.50e-06 ***
sexFemale    -0.360725   0.072800  -4.955 7.23e-07 ***
age           0.019221   0.004614   4.165 3.11e-05 ***
smokeCurrent -0.927517   0.098896  -9.379 < 2e-16 ***
smokeFormer  -0.179407   0.081154  -2.211 0.0271 *
htnYes       -1.541731   0.082703 -18.642 < 2e-16 ***
hypercholYes -1.345193   0.065272 -20.609 < 2e-16 ***
famhistYes   -0.951521   0.070274 -13.540 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8750.7 on 6323 degrees of freedom
Residual deviance: 7589.2 on 6316 degrees of freedom
AIC: 7605.2

Number of Fisher Scoring iterations: 4
```

○ En aquest tercer model es poden identificar aquelles variables amb associació significativa respecte a la variable diabetis:

- OR d'aquestes variables en relació a la diabetis:

		2.5 %	97.5 %
(Intercept)	4.37	2.29	8.34
sexFemale	0.70	0.60	0.80
age	1.02	1.01	1.03
smokeCurrent	0.40	0.33	0.48
smokeFormer	0.84	0.71	0.98
htnYes	0.21	0.18	0.25
hypercholYes	0.26	0.23	0.30
famhistYes	0.39	0.34	0.44

- Atenent a la taula d'OR presentada, aquestes dades ens permet veure l'associació entre les variables definides i la variable d'interès: la diabetis:
 - Sexe: Sembla que hi ha una associació menor entre la diabetis i el gènere femení.
 - Edat: l'increment de l'edat es pot veure associat amb major presència de diabetis.
 - Per a la resta de dades (estat de l'hàbit de consum de tabac, hipertensió, hipercolesterolèmia i història familiar de patologia prematura cardiovascular) sembla que la relació amb la diabetis és inversa. Aquesta relació pot tenir origen en protocols d'atenció específics a les persones amb diabetis que modifiquin l'aparició de dites variables o, inclòs, pot haver-hi un efecte de la distribució de la població en els grups tal com es manifesta en la següent taula:

-----Summary descriptives table by 'Type-2 diabetes'-----

	No N=3322	Yes N=3002	p. overall
Sex:			<0.001
Male	1267 (38.1%)	1412 (47.0%)	
Female	2055 (61.9%)	1590 (53.0%)	
Age	66.6 (6.10)	67.5 (6.22)	<0.001
Smoking:			<0.001
Never	2053 (61.8%)	1839 (61.3%)	
Current	514 (15.5%)	344 (11.5%)	
Former	755 (22.7%)	819 (27.3%)	
Body mass index	30.0 (3.63)	29.9 (4.01)	0.220
Waist circumference	99.6 (10.6)	101 (10.5)	<0.001
Hypertension:			<0.001
No	283 (8.52%)	806 (26.8%)	
Yes	3039 (91.5%)	2196 (73.2%)	
Dyslipidemia:			<0.001
No	524 (15.8%)	1222 (40.7%)	
Yes	2798 (84.2%)	1780 (59.3%)	
Family history of premature CHD:			<0.001
No	2366 (71.2%)	2529 (84.2%)	
Yes	956 (28.8%)	473 (15.8%)	

- En aquesta taula de comparació es pot observar que la distribució de les persones amb o sense diabetis presenta diferències significatives per a les diverses variables.
- Aquest element pot causar una distribució desigual que modifiqui l'associació real entre les variables seleccionades i la variable objecte (en aquest cas la diabetis).

La adherència a la dieta mediterrània:

L'adherència a la dieta mediterrània es posiciona com una variable o factor que pot resultar protector davant les malalties agudes o degeneratives derivades de l'alimentació i que presenten una afectació cardíoc-vascular. En aquest sentit, ja s'ha pogut observar que aquesta variable presenta una OR de 0,95 amb un IC: 0,93-0,98.

En l'estudi Predimed tracten de buscar un mecanisme per a poder valorar l'adherència que presenten els diferents participants de l'estudi a la dieta mediterrània a partir d'un recull d'informació centrat en un test que permet detectar i determinar els hàbits de consum dels participants per a poder construir un nexa la habitualitat i fixació que presenten aquests a la dieta amb aliments propis de la regió mediterrània com a preferència de consum. Aquest element és quelcom que els investigadors consulten des de la primera sessió de trobades amb els participants: en aquest sentit, s'aplica el test en la primera entrevista i es va repetint en les trobades trimestrals.

Atenent a aquesta circumstància, es proposa una anàlisi on es relaciona la variable relativa a la puntuació de l'adherència (p14): estratificada per grups de puntuació (0,4,6,8,10,12,14) en relació a l'aparició o absència de l'esdeveniment objecte d'aquestes estudi.

-----Summary descriptives table by 'AMI, stroke, or CV Death'-----

	No N=6062	Yes N=251	OR	p.ratio	p.overall
Gp14:					
(12,14]	91 (1.50%)	3 (1.20%)	Ref.	Ref.	
(0,4]	99 (1.63%)	8 (3.19%)	2.37 [0.65;11.7]	0.201	
(4,6]	665 (11.0%)	34 (13.5%)	1.48 [0.52;6.48]	0.504	
(6,8]	1951 (32.2%)	92 (36.7%)	1.36 [0.50;5.81]	0.590	
(8,10]	2302 (38.0%)	91 (36.3%)	1.14 [0.42;4.87]	0.821	
(10,12]	954 (15.7%)	23 (9.16%)	0.70 [0.24;3.12]	0.591	

D'aquesta taula es poden observar les següents qüestions:

- Es pot observar que, en totes les puntuacions d'adherència a la dieta mediterrània, la majoria d'interval de confiança de les OR inclouen el valor d'igualtat de risc (1). Aquest fet ens permet determinar, que en tots els casos, encara que els riscos varien de forma inversa a la puntuació obtinguda al test, no ens permet concloure que l'adherència a la dieta sigui un factor determinant per a la prevenció o augment del risc de malalties o mort per causes cardíoc-vasculars.

En quant a les taules de contingència especificades a continuació, es fa una anàlisi amb l'enfocament que entén que els participants disposaven d'una adherència a la dieta mediterrània determinada anteriorment a l'inici de l'estudi: i que a partir de l'estudi es comença a recollir com a dada. Atenent a aquesta orientació, es pot analitzar que, certes casuístiques cronicades com la diabetis, hipertensió o hipercolesterolèmia poden tenir relació amb, a llarg termini, amb l'adherència a la dieta mediterrània que poguessin disposar els participants abans d'iniciar l'estudi. D'aquestes taules es pot determinar:

- En relació a la hipertensió:

-----Summary descriptives table by 'Hypertension'-----

	No N=1089	Yes N=5224	OR	p.ratio	p.overall
Gp14:					0.186
(12,14]	24 (2.20%)	70 (1.34%)	Ref.	Ref.	
(0,4]	19 (1.74%)	88 (1.68%)	1.58 [0.80;3.16]	0.186	
(4,6]	116 (10.7%)	583 (11.2%)	1.73 [1.02;2.83]	0.041	
(6,8]	332 (30.5%)	1711 (32.8%)	1.77 [1.08;2.82]	0.025	
(8,10]	416 (38.2%)	1977 (37.8%)	1.64 [1.00;2.60]	0.052	
(10,12]	182 (16.7%)	795 (15.2%)	1.50 [0.90;2.43]	0.115	

- En termes generals s'observa un relació poc significativa entre la variable d'adherència a la dieta mediterrània respecte a la hipertensió.
- En valors mitjos d'adherència a la dieta mediterrània s'observa una relació de risc respecte a la possibilitat de patir hipertensió per no seguir aquesta determinada dieta. En aquest sentit, per a persones amb adherència de 4 a 6 punts i de 6 a 8 punts, els OR i IC són, respectivament, de: 1.73 [1.02,2.83] i 1.77 [1.08-2.82].

- En relació a la dislipèmia:

-----Summary descriptives table by 'Dyslipidemia'-----

	No N=1745	Yes N=4568	OR	p.ratio	p.overall
Gp14:					0.808
(12,14]	24 (1.38%)	70 (1.53%)	Ref.	Ref.	
(0,4]	26 (1.49%)	81 (1.77%)	1.07 [0.56;2.04]	0.841	
(4,6]	198 (11.3%)	501 (11.0%)	0.87 [0.52;1.41]	0.580	
(6,8]	564 (32.3%)	1479 (32.4%)	0.90 [0.55;1.43]	0.672	
(8,10]	676 (38.7%)	1717 (37.6%)	0.87 [0.53;1.38]	0.576	
(10,12]	257 (14.7%)	720 (15.8%)	0.96 [0.58;1.55]	0.884	

- En tots els casos, els intervals de confiança per a les OR inclouen el valor 1 i no permet determinar un risc major de hipercolesterolèmia segons el nivell d'adherència a la dieta mediterrània.

- En relació a la diabetis:

-----Summary descriptives table by 'Type-2 diabetes'-----

	No N=3317	Yes N=2996	OR	p.ratio	p.overall
Gp14:					0.057
(12,14]	52 (1.57%)	42 (1.40%)	Ref.	Ref.	
(0,4]	53 (1.60%)	54 (1.80%)	1.26 [0.72;2.21]	0.417	
(4,6]	362 (10.9%)	337 (11.2%)	1.15 [0.75;1.78]	0.523	
(6,8]	1036 (31.2%)	1007 (33.6%)	1.20 [0.79;1.83]	0.385	
(8,10]	1259 (38.0%)	1134 (37.9%)	1.11 [0.74;1.70]	0.609	
(10,12]	555 (16.7%)	422 (14.1%)	0.94 [0.61;1.45]	0.780	

- En tots els casos, els intervals de confiança per a les OR inclouen el valor 1 i no permet determinar un risc major de patir diabetis segons el nivell d'adherència a la dieta mediterrània.

ANÀLISI DE SUPERVIVÈNCIA

En aquest apartat tenim com a objectiu processar les dades on l'aspecte important serà el temps que transcorre fins que apareixen les complicacions cardiovasculars. És una situació molt important en estudis biomèdics, per exemple per avaluar factors de risc en l'avaluació temporal de l'aparició de diferents problemes, com la mateixa mortalitat, d'aquí el nom, fins l'aparició de malalties.

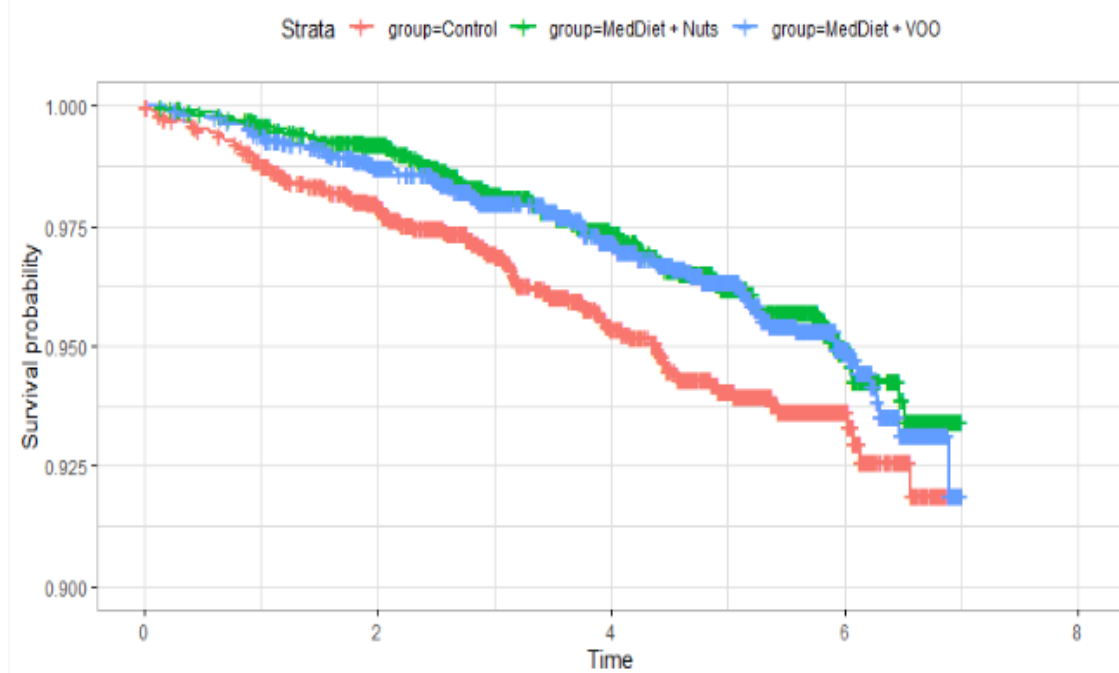
Si ens fixem amb el gràfic de supervivència que es presenta a continuació, podem observar els resultats que relatius a la evolució de la supervivència i comparar-les entre els grups contemplats dins l'estudi.

Com es pot observar, trobem representat el grup control de color rosat, tenint la probabilitat més baixa.

Seguidament, de color blau tenim representat el grup que se li va administrar una dieta rica en oli d'oliva verge extra, tenint una supervivència una mica més alta que el grup control.

I finalment, de color verd trobem el grup que va prendre una dieta rica en fruits secs, amb la probabilitat de supervivència més alta.

Anteriorment, en la significació de variables, ja s'havia vist com la dieta rica en fruits secs tenia un p-valor més significatiu (0,02) que no pas en el cas de la dieta rica en oli d'oliva verge extra (0,17).



Per a poder considerar l'efecte de diverses variables respecte a la supervivència, es presenta una anàlisi basat en la regressió de Cox. A través d'aquesta tipologia analítica es podrà comparar grups i interpretar els factors de risc.

En primer lloc, es proposa una anàlisi centrada en aquest mètode relacionant la supervivència al tipus de dieta o grups inclosos dins l'estudi:


```

Call:
coxph(formula = pred.surv ~ group, data = predimednew)

n= 6324, number of events= 252

              coef exp(coef) se(coef)      z Pr(>|z|)
groupMedDiet + Nuts -0.4190   0.6577  0.1569 -2.671  0.00757 **
groupMedDiet + VOO -0.3526   0.7028  0.1487 -2.371  0.01775 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
groupMedDiet + Nuts   0.6577     1.520   0.4836   0.8945
groupMedDiet + VOO   0.7028     1.423   0.5251   0.9407

Concordance= 0.559 (se = 0.018 )
Likelihood ratio test= 8.62 on 2 df,  p=0.01
Wald test               = 8.89 on 2 df,  p=0.01
Score (logrank) test = 9 on 2 df,  p=0.01

```

Tenint en compte les dades que ens aporten els exponencials del coeficient, podem observar que, tant la dieta mediterrània amb suplementació d'OOVE com la mateixa dieta amb suplementació a partir de fruits secs, disminueixen el rati de risc:

- En el cas de dieta mediterrània + OOVE: l'exponencial és de 0.70 amb un IC d'entre 0.52 i 0.94.
- Aquesta reducció és major en el cas de la dieta mediterrània + fruits secs on s'observa un exponencial de 0.65 amb un IC d'entre 0.48 i 0.89.

A continuació es proposa un nou model per a la regressió de Cox tenint en compte les diverses variables contemplades per l'estudi, per a veure la seva implicació en la supervivència.

Model 1:

```

Call:
coxph(formula = pred.surv ~ group + sex + age + smoke + bmi +
      waist + wth + htn + diab + hyperchol + famhist + hormo +
      p14, data = predimednew)

n= 5661, number of events= 224
(663 observations deleted due to missingness)

              coef exp(coef) se(coef)      z Pr(>|z|)
groupMedDiet + Nuts -0.46983   0.62511  0.16933 -2.775  0.005526 **
groupMedDiet + VOO -0.36749   0.69247  0.15748 -2.333  0.019622 *
sexFemale           -0.70082   0.49618  0.22732 -3.083  0.002049 **
age                 0.06292   1.06494  0.01150  5.469  4.51e-08 ***
smokeCurrent        0.54649   1.72718  0.22679  2.410  0.015965 *
smokeFormer         0.45103   1.56993  0.19306  2.336  0.019480 *
bmi                 -0.03790   0.96281  0.02889 -1.312  0.189514
waist               -0.01956   0.98063  0.01786 -1.096  0.273275
wth                 5.99650  402.02126  2.93267  2.045  0.040882 *
htnYes              0.47122   1.60195  0.19784  2.382  0.017229 *
diabYes             0.54723   1.72846  0.14783  3.702  0.000214 ***
hypercholYes       -0.24320   0.78412  0.14095 -1.725  0.084449 .
famhistYes          0.24261   1.27457  0.16808  1.443  0.148907
hormoYes            -0.85333   0.42599  1.00724 -0.847  0.396885
p14                 -0.10455   0.90073  0.03351 -3.120  0.001810 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
groupMedDiet + Nuts   0.6251   1.599723   0.44856  8.711e-01
groupMedDiet + VOO   0.6925   1.444100   0.50857  9.429e-01
sexFemale             0.4962   2.015408   0.31779  7.747e-01
age                   1.0649   0.939023   1.04120  1.089e+00
smokeCurrent          1.7272   0.578979   1.10738  2.694e+00
smokeFormer           1.5699   0.636970   1.07534  2.292e+00
bmi                   0.9628   1.038631   0.90980  1.019e+00
waist                 0.9806   1.019756   0.94690  1.016e+00
wth                   402.0213  0.002487   1.28221  1.260e+05
htnYes                1.6019   0.624240   1.08703  2.361e+00
diabYes               1.7285   0.578551   1.29369  2.309e+00
hypercholYes          0.7841   1.275322   0.59485  1.034e+00
famhistYes            1.2746   0.784581   0.91684  1.772e+00
hormoYes              0.4260   2.347456   0.05916  3.067e+00
p14                   0.9007   1.110216   0.84346  9.619e-01

Concordance= 0.714 (se = 0.016 )
Likelihood ratio test= 134.2 on 15 df,  p=<2e-16
Wald test               = 132.5 on 15 df,  p=<2e-16
Score (logrank) test = 137.8 on 15 df,  p=<2e-16

```

En aquest primer model es contemplen totes les variables per a poder determinar la seva incidència en la supervivència. Arran d'aquesta anàlisi es pot detectar que:

- Les variables grup, sexe, edat, smoke, rati maluc-alçada, hipertensió, diabetis i el test d'adherència a la dieta mediterrània presenten una implicació significativa respecte a la supervivència:
 - Les que hi tenen una implicació protectora podrien ser: la dieta mediterrània amb fruits secs (0.62. IC: 0.44-0.87), dieta mediterrània amb OOVE (0.69. IC: 0.51-0.94), sexe femení (0.49. IC: 0.31-0.77) i p14 (0.90. IC: 0.84-0.96)
 - Les que hi tenen una implicació de risc serien: hàbit de fumar actual (1.72, IC: 1.10- 2.69), hàbit de fumar abandonat (1.56. IC: 1.07-2,29), rati alçada-maluc (402. IC: 1.28-126000), hipertensió (1.60. IC: 1.08-2.36) i diabetis (1.72. IC: 1.29-2.30).
- Les variables no significatives detectades en el model són bmi, waist, hyperchol, famhist i homo. Es proposa un nou model descartant aquestes variables.

Model 2:

```
Call:
coxph(formula = pred.surv ~ group + sex + age + smoke + wth +
      htn + diab + p14, data = predimednew)

n= 6324, number of events= 252

              coef exp(coef) se(coef)      z Pr(>|z|)
groupMedDiet + Nuts -0.37403  0.68796  0.15786 -2.369 0.017817 *
groupMedDiet + VOO -0.32203  0.72468  0.14928 -2.157 0.030988 *
sexFemale           -0.50013  0.60645  0.17392 -2.876 0.004031 **
age                 0.06731  1.06963  0.01031  6.526 6.75e-11 ***
smokeCurrent        0.69383  2.00137  0.20700  3.352 0.000803 ***
smokeFormer         0.50002  1.64875  0.18036  2.772 0.005565 **
wth                 1.34616  3.84263  1.00230  1.343 0.179248
htnYes              0.30787  1.36052  0.17547  1.755 0.079338 .
diabYes             0.58064  1.78718  0.13628  4.261 2.04e-05 ***
p14                 -0.10998  0.89585  0.03099 -3.548 0.000388 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
groupMedDiet + Nuts    0.6880    1.4536    0.5049    0.9374
groupMedDiet + VOO    0.7247    1.3799    0.5408    0.9710
sexFemale              0.6064    1.6489    0.4313    0.8528
age                   1.0696    0.9349    1.0482    1.0915
smokeCurrent           2.0014    0.4997    1.3339    3.0028
smokeFormer            1.6488    0.6065    1.1578    2.3479
wth                   3.8426    0.2602    0.5389   27.4021
htnYes                 1.3605    0.7350    0.9646    1.9190
diabYes                1.7872    0.5595    1.3683    2.3343
p14                    0.8959    1.1163    0.8431    0.9520

Concordance= 0.705 (se = 0.016 )
Likelihood ratio test= 135.7 on 10 df,  p=<2e-16
Wald test               = 135.9 on 10 df,  p=<2e-16
Score (logrank) test = 140.1 on 10 df,  p=<2e-16
```

D'aquesta anàlisi es pot observar un model on:

- Les variables grup, sexe, edat, hàbit de fumar, diabetis i p14 romanen com a significatives.
 - Les que hi tenen una implicació protectora podrien ser: la dieta mediterrània amb fruits secs (0.68. IC: 0.50-0.93), dieta mediterrània OOVE (0.72. IC: 0.54- 0.97), sexe femení (0.60. IC: 0.43-0.85) i p14 (0.89. IC:0.84-0.95).

- Les que hi tenen una implicació de risc serien: hàbit de fumar actual (2.00, IC: 1.33- 3.00), hàbit de fumar abandonat (1.64. IC: 1.15-2,34) i diabetis (1.78. IC: 1.36-2.33).
- Les variables d'hipertensió i rati alçada-maluc perden significança. Es proposa un model descartant aquestes darreres variables.

Model 3:

```
Call:
coxph(formula = pred.surv ~ group + sex + age + smoke + diab +
      p14, data = predimednew)

n= 6324, number of events= 252

              coef exp(coef) se(coef)      z Pr(>|z|)
groupMedDiet + Nuts -0.39103  0.67636  0.15764 -2.480 0.013121 *
groupMedDiet + VOO  -0.33799  0.71320  0.14907 -2.267 0.023372 *
sexFemale           -0.45051  0.63730  0.17037 -2.644 0.008184 **
age                 0.06924  1.07170  0.01026  6.749 1.49e-11 ***
smokeCurrent        0.66955  1.95336  0.20533  3.261 0.001111 **
smokeFormer         0.50756  1.66123  0.17886  2.838 0.004544 **
diabYes             0.52958  1.69821  0.13299  3.982 6.83e-05 ***
p14                 -0.11587  0.89059  0.03072 -3.772 0.000162 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              exp(coef) exp(-coef) lower .95 upper .95
groupMedDiet + Nuts    0.6764    1.4785    0.4966    0.9212
groupMedDiet + VOO     0.7132    1.4021    0.5325    0.9552
sexFemale              0.6373    1.5691    0.4564    0.8899
age                   1.0717    0.9331    1.0504    1.0935
smokeCurrent           1.9534    0.5119    1.3062    2.9212
smokeFormer            1.6612    0.6020    1.1700    2.3587
diabYes                1.6982    0.5889    1.3085    2.2039
p14                   0.8906    1.1228    0.8386    0.9459

Concordance= 0.7 (se = 0.016 )
Likelihood ratio test= 130.3 on 8 df,  p=<2e-16
Wald test              = 131 on 8 df,  p=<2e-16
Score (logrank) test = 135 on 8 df,  p=<2e-16
```

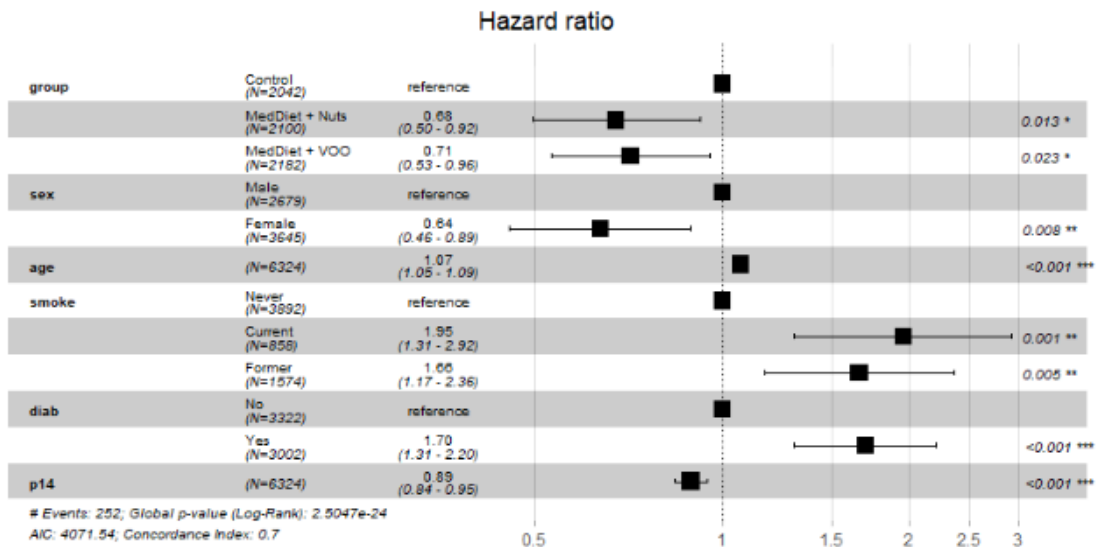
- En aquest darrer model, totes les variables referenciades presenten un p-valor significatiu en quant a la seva implicació en la supervivència. Amb aquest model podem concloure que:
 - Les que hi tenen una implicació protectora podrien ser: la dieta mediterrània amb fruits secs (0.67. IC: 0.50-0.92), dieta mediterrània OOVE (0.71. IC: 0.53- 0.95), sexe femení (0.63. IC: 0.46-0.89) i p14 (0.89. IC:0.84-0.95).
 - Les que hi tenen una implicació de risc serien: hàbit de fumar actual (1.95, IC: 1.30- 2.92), hàbit de fumar abandonat (1.66. IC: 1.17-2,35) i diabetis (1.70. IC: 1.30-2.20).

No obstant això, s'aplica una anàlisi de la variància dels models per a detectar aquell model amb més significança:

```
Analysis of Deviance Table
Cox model: response is pred.surv
Model 1: ~ group
Model 2: ~ group + sex + age + smoke + diab + p14
Model 3: ~ group + sex + age + smoke + wth + htn + diab + p14
  loglik   Chisq Df P(>|Chi|)
1 -2088.6
2 -2027.8 121.6376  6 < 2e-16 ***
3 -2025.1  5.3967  2  0.06732 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- A través d'aquesta anova es pot observar una millora del model 2 respecte al primer.
- Per aquest motiu, es proposa tenir en compte el segon model.

En relació a aquest segon model, es presenta a continuació una gràfica sobre els ratis de risc per a cada variable:



De la gràfica anterior es pot extreure les següents conclusions:

- Es pot observar el caràcter protector que presenten el seguiment de dietes mediterrànies amb suplementació de greixos saludables (com OOVE i fruits secs) en contrast amb la dieta baixa en greixos (control).
- La variable del sexe ens permet observar una supervivència major en el sexe femení.
- Per edat, cada any que es pugui augmentar fa que hi hagi un risc major de patir malalties cardíco-vasculars.
- Els hàbits de fumar (sigui actual o passat) augmenten el risc de patir, en el futur, les malalties cardíco-vasculars proposades com a end-point en l'estudi. En aquest cas, s'observa un risc major si el consum de tabac persevera.
- La diabetis suposa una condició que fa augmentar el risc de patir aquesta tipologia de malalties.
- Per acabar, sembla que l'adherència a la dieta mediterrània, tal com s'ha recollit en l'estudi, pot tenir una implicació protectora davant el risc de patir aquestes malalties.

CONCLUSIONS:

Després de realitzar aquest exercici analític es pot confirmar la idea general que presenta l'estudi PREDIMED, el qual consisteix en afirmar que una dieta mediterrània rica en greixos saludables, concretament rica en OOVE o bé en fruits secs, pot ajudar a prevenir malalties cardiovasculars, i com a conseqüència directa, prevenir la mort que pot provocar aquest tipus de malalties.

Si ens centrem en les dues dietes mediterrànies que s'utilitzen al llarg de l'estudi, s'ha pogut observar com una dieta rica en fruits secs té un major efecte preventiu que en la dieta amb OOVE sobre la casuística de patir problemes cardiovasculars.

Tot i així és molt rellevant no deixar de banda altres factors ja que també s'ha pogut veure que tenen molta importància a l'hora de patir aquest tipus de patologies. Per exemple, el sexe femení o tenir una alta adherència a la dieta, junt a l'alimentació que segueixen els participants de l'estudi, pot ajudar a prevenir aquest tipus de malalties. D'altra banda, fumar o patir diabetis en pot incrementar el risc.

Pel que fa a la anàlisi de la diabetis, cal puntualitzar que s'ha dissenyat dins el marc d'un estudi amb uns objectius d'investigació que no s'han centrat en descriure ni investigar aquesta patologia específicament. Aquest fet comporta que els resultats obtinguts permetin obtenir una associació entre la variable diabetis i altres variables contemplades en l'estudi sense poder establir una relació de causalitat. Per tal fet, es consideraria oportú investigar aquesta associació amb un disseny o l'end-point fos, pròpiament, la patologia de la diabetis.

En relació a les dietes mediterrànies i alguns dels factors que augmenten el risc de patir malalties cardiovasculars o patologies cròniques o de morbiditat associades. En aquest sentit, no s'ha pogut concloure que una major adherència a aquestes dietes pugui reduir el risc de patir algun d'aquests factors.

Cal destacar que algunes de les condicions que s'utilitzaven a l'hora d'escollir els subjectes per realitzar l'estudi no tenen rellevància a l'hora de reduir o augmentar el risc de patir malalties cardiovasculars.

Per tant, si s'hagués de repetir l'estudi de Prevenció amb dieta mediterrània, seria adequat eliminar totes aquelles condicions que s'ha vist que no intervenen sobre els resultats finals i se'n podrien afegir de noves per tal d'observar els efectes o implicacions que poden tenir alguns factors que en aquest estudi no s'han tingut en compte: per exemple l'activitat física o laboral, el consum d'alcohol, i si ens fixem en l'actualitat, haver patit el COVID-19. També es podria afegir un altre tipus de dieta, com per exemple una dieta rica en peix blau com el salmó, i veure quins efectes final pot tenir.

Per acabar, tot i que la mostra que s'utilitza és bastant elevada, es podria augmentar aquesta mostra i d'aquesta manera veure si s'obtenen els mateixos resultats i si es poden extreure conclusions que en aquest cas no s'han pogut afirmar. Paral·lelament, seria oportú poder aplicar aquesta tipologia d'estudi amb participants que no tinguin un alt risc de patir malalties cardiovasculars per tal de veure l'efecte, a llarg termini, d'aquesta tipologia de dietes.