Author's response to reviews

Title: Health-related quality of life in epilepsy adults: the effect of age-related factors in a multicentre Italian study

Authors:

Valeria Edefonti (valeria.edefonti@unimi.it)
Francesca Bravi (francesca.bravi@unimi.it)
Katherine Turner (katherine.turner@ao-sanpaolo.it)
Ettore Beghi (ettore.beghi@marionegri.it)
Maria Paola Canevini (maria.canevini@ao-sanpaolo.it)
Monica Ferraroni (monica.ferraroni@unimi.it)
Ada Piazzini (ada.piazzini@ao-sanpaolo.it)

Version: 5 Date: 16 February 2011

Author's response to reviews: see over
Dear Editor-in-Chief,

following the letter written on February, 2nd, 2011, thank you for considering our manuscript, “Health-related quality of life in epilepsy adults: the effect of age-related factors in a multicentre Italian study”.

We are enclosing the second revision of the manuscript after the modifications suggested by referee 2. Please find our point-by-point responses to the reviewer’s comments below (Q for question and A for answer). Briefly, we accepted most of his suggestions, but we decided to simplify statistical comments in the text, instead of putting them in a separate appendix. We strongly believe that it is important to provide statistical details in the text in the hope that this improves confidence with statistics in the medical community.

I would be happy to assist your editorial staff in the event that further questions arise, and I look forward to receiving your final decision.

Thank you for your attention.

Yours sincerely,

Valeria Edefonti, Ph.D.
Sezione di Statistica Medica e Biometria “Giulio A. Maccacaro”
Dipartimento di Medicina del Lavoro “Clinica del Lavoro L. Devoto”
Via Venezian, 1
20133 Milano, Italy
Tel: +39-02-50320873; Fax: +39-02-50320866
E-mail valeria.edefonti@unimi.it
Reviewer's report:
The authors have improved this paper significantly but, in the opinion of the reviewer, there are still major reservations. Some of these may be rectifiable.
Q1. The abstract remains unclear.
A1. We modified the abstract according to your suggestions (see the next points).

Q2. The study was carried out on a selected sample. There is no control group. The patient group is a selected one. The authors have not taken this adequately into account in their discussion. These factors could affect the results greatly.
A2. We added this information in the abstract: “We examined this association in a selected population of 815 adults with epilepsy recruited in the context of a multicentre study for the evaluation of Epi-QoL.” However, the study was based on a specific instrument that measures health-related quality of life for patients with epilepsy, so, by definition, there could not be any control group filling in the same questionnaire.

Q3. If the reviewer has understood the situation correctly, the authors have embarked on carrying out an analysis which could be stated quite simply as follows. “Health-related quality of life has been assessed in a selected group of adults with epilepsy, examining three age-related factors: age at onset, duration of epilepsy and age of the patient. Because these parameters are inter-related, the statistical analysis has included not only an examination of each of these factors on its own but also a regression analysis examining each of the factors while controlling for the other two.” The authors could then simply state what the results were, according to these types of analyses.
A3. We do not think this synthesis works. We fitted regression models for one age-related factor at a time (6 models, 3 of which including the selected confounding variables) and for pairs of age-related factors (6 models, 3 of which including the selected confounding variables). We did not fit a regression model including the three age-related factors simultaneously, because they are correlated. Anyway, we modified the sentences on statistical analysis in the abstract and through the text (see below).

Q4. If the authors are of the opinion that details of the statistical analysis should appear then this information could be provided in an appendix, leaving the reader a much simpler task of examining the study outlined in the reviewer’s foregoing comments (in quotation marks above). The body of the paper would then consist of how the patients were recruited, a list of the hypothesis tested, the results obtained and a detailed discussion of the results, including a critical review of how these agree or disagree with previous studies, followed by appropriate conclusions.
A4. We simplified single sentences providing statistical details, but we strongly believe that this part should be left in the text. The other referees agree with this idea too.

Some detailed comments follow.
Comments that would probably not be comprehensible to the non-statistician reader will be preceded by the words “Statistical comment” in what follows. These should either be rewritten in language that the general scientist could
understand or should be moved to an appendix.

Page 1.
Q5. The usual word is “multicentre” not “multicentric”.
A5. We modified the text accordingly.

Q6. Statistical comment. “We fitted simple regression models including age-related factor alone. We also fitted multiple regression models including pairs of age-related factors solely, as well as one or two age-related factors together with the same set of confounders.”
A6. We modified the text in the following way: “Ordinary least-squares regression models were used to assess the relationships between age-related factors (patient’s age, age at seizure onset, and duration of epilepsy) and overall Epi-QoL score, controlling for the effect of potential confounders. We fitted simple regression models including each age-related factor alone to assess the independent role of each factor on the overall Epi-QoL score. We also fitted multiple regression models including pairs of age-related factors solely, as well as one or two age-related factors together with the same set of confounders.”.

Q7. Statistical comment. “Multiple regression models including two age-related factors show that duration of epilepsy is still a significant predictor of the overall Epi-QoL score in both pairwise models, whereas age is a significant negative only in the model including age at onset. Age at onset emerges as a significant positive predictor of the overall Epi-QoL score only in the model including age: the higher is age at onset, the higher is the overall Epi-QoL score. Adjusted regression models including one or two age-related factors, as well as the selected confounding variables, show that the age-related factors had no significant effect on the overall Epi-QoL score.” With regard to the last sentence, this appears to contradict the conclusion that age-related factors are important. The authors should explain this apparent contradiction simply and clearly both in the abstract and the body of the paper. The authors need to be more precise in the last line of the abstract. The statement is as follows.
“However, additional information on demographic and clinical factors may provide a better explanation of HRQOL in epilepsy.”
A7. We modified the text in the following way: “Adjusted regression models including either one or two age-related factors and controlling for the selected confounding variables show that the age-related factors have no significant effect on the overall Epi-QoL anymore. If no other known correlates of the overall Epi-QoL score are considered, age and duration of epilepsy can be expected to have a significant negative association with HRQOL in epilepsy (with the effect of duration being stronger and more consistent across models than the one of age), whereas age at onset is a positive predictor of the overall HRQOL of limited significance. However, demographic and clinical factors, such as seizure frequency in the preceding 12 months, may provide a better explanation of HRQOL in epilepsy.”.

Page 2 (continued).
Q8. Replace “each of the previous factors” with “each of these factors”.
A8. We modified the text accordingly.
Q9. The following sentence needs to be rewritten both from the point of view of language and also to enable the reader to understand what the authors intended. “The simultaneous evaluation of two of these factors allows to take into account existing compounding phenomena.”

The reviewer did not understand the argument that the authors were trying to put forward here. Would it not be more appropriate to suggest that each of the factors was examined with and without controlling for the others?

A9. We modified the text in the following way: “Including two age-related factors in the same model allows to estimate the effect of interest adjusting for the effect of the remaining age-related factor, and this may provide results that are different from the ones of simple models including only the same age-related factor of interest, in consideration of confounding phenomena.”

Page 3.

Q10. Statistical comment. “The major implications of this issue is that the overall picture on age-related morbidity and HRQOL is only provided by the joint evaluation of the three models including single pairs of age-related factors.”

Again, the reviewer did not understand this statement. Could the authors not simply state that they avoided analysing any two related factors at the same time, indicating simply how they did this.

A10. No, this is not true, as we did it. There is no way to put the three age-related factors in the same model. Anyway, we modified the text in the following way: “The major implication of this issue is that one can get a good overall picture on age-related morbidity and HRQOL only from the joint evaluation of the three models including each available pair of age-related factors.”

Q11. The next sentence finishes with the words “compared with other disease-related factors”. The authors have not made clear what they are comparing with what.

A11. We deleted this part of the sentence.

Q12. Where the authors have written “We identified one study” do they mean “We identified only one study”? A12. We modified the text accordingly.

Q13. Are the following words necessary or could they be deleted? “Including generic and specific questions about the more relevant determinations of HRQOL measured according to a”? The sentence could then simply finish “using a variety of validated instruments…”

A13. We simplified the text as suggested.

Page 4.

Q14. Statistical comment. The section beginning statistical analyses up to and including the heading results should, in the opinion of the reviewer, all be placed in an appendix and should be replaced with a simple explanation of the essentials of the statistical analysis that would be comprehensible to a non-statistician scientist.

A14. We do not agree on this suggestion and just try to simplify the statistical comments provided in the body of the paper. For instance, we modified the sentence on kernel density estimation in the following way: “We computed a
kernel density estimate of the distribution of each age-related factor and of the overall Epi-QoL score to improve histogram-like representations of the original data.”.

Page 5 (continued).
Q15. Replace “remission since more than a year” with “remission for more than a year”.
A15. We modified the text as suggested.

Q16. Replace “all showed non-drug-resistance seizures” with “all had non-drug-resistant seizures”.
A16. We modified the text as suggested.

Q17. If the last two sentences on this page are to be included in the body of the paper the authors should explain what the kernel densities are in language that a non-statistician scientist can understand.
A17. We modified the sentences on kernel densities in both Statistical Analysis and Results Sections in the following way:
Statistical Analysis:
“We computed a kernel density estimate of the distribution of each age-related factor and of the overall Epi-QoL score to improve histogram-like representations of the original data.”
Results:
“The density estimates provide an indication of the general tendency of each variable to distribute on the range of its possible values. Here, all the density estimates emerged to be highly skewed, with the age-related densities that were skewed to the right and the overall Epi-QoL score density that was skewed to the left.”

Page 6.
Q18. In the sentence beginning: “Age was positively correlated with age at onset...”, the authors are referring to relationships between three variables that would be inter-related in any case. It was difficult to see what the point of this was.
A18. The calculation of correlation coefficients allows to measure direction and strength of the linear relationship that potentially links two variables in a given sample. So, you get an information on pairs of variables that you suppose to be correlated in theory, but on which you do not know anything in practice (ie in your sample of patients). And, in any way, you do not know anything about the sign of their correlation coefficient.

Q19. In the next sentence beginning: “The overall Epi-QoL score was correlated...”, with the potential confounders “psychiatric disturbances, number of drugs and intellectual functions” the authors do not provide details of these very broad variables nor do they give information on either the degree or the significance of the correlations.
A19. We added a reference for details on these variables at the end of the sentence in the following way: “(see [34] and the beginning of this section for details on these variables)”.
Moreover, Table 1 already showed statistically significant correlation coefficients in bold typeface. We did not report in the text correlation coefficients and
corresponding p-values because this was not the main focus of the paper.

Q20. Since, two sentences later, they state that in the simple regression models age at onset was not a significant predictor, this issue should be clarified better in the discussion.

A20. We clarified this issue in the Discussion in two points: “It emerges that, if no other known correlates of the overall Epi-QoL score are considered, age and duration of epilepsy can be expected to have a significant consistent negative association, and age at onset a positive association of limited significance (i.e. significant in one model only), with HRQOL in epilepsy. However, none of the age-related variables showed a significant effect on the overall Epi-QoL score, after adjusting for a selected set of confounders including demographic and clinical factors, and, in particular, the seizure frequency in the preceding 12 months.” and “Finally, in our paper, age at onset emerged as a positive predictor of limited significance (in one model out of three) for the overall Epi-QoL.”.

We also added this information in the abstract.

Q21. Replace “to the one emerged” with “to the one that emerged”.

A21. We modified the text as suggested.

Q22. In the second half of this page the authors use the word “including” or “included” at least five times. Do they mean that the analysis was “controlled for” these factors or what do they mean?

A22. A regression model “include” as independent variables one or more variables of interest and some adjusting variables. All of them are “included” in the model, but have a different interpretation. Age-related factors are always variables of interest and the confounders are always adjusting variables to control for, but the effect of one age-related factor may be adjusted for the effect of another one or not. Moreover, a simple regression model include only one independent variable and, by definition, do not include variables to adjust for. So, “include” is just a general term that fits for all the different situations.

Q23. Where the authors have written “to having less than 30 years” do they mean “to being less than 30 years”?

A23. Yes we do, and we modified the text as suggested.

Q24. The authors should distinguish more clearly between what a type-1 model and a type-2 model is, for the non-statistician reader.

A24. We modified the sentences in the text as follows: “Type-1 models include only the age-related variables (either one or two of them) as predictors of the overall Epi-QoL score. Type-2 models include the age-related variables (either one or two of them) and the selected potential confounders (as listed in the left outer column of the tables).” We added a footnote on type-1 and type-2 models in Tables 2 and 3.

Q25. The second-last sentence before the discussion would be incomprehensible to the non-statistician reader. The authors should explain this in plain English.

A25. We modified the sentence adding an explanation on what $R^2$ is and referring to type-1 and type-2 models to avoid confusion with the “adjusting $R^2$“.
sentence is now as follows: “This is evident in the comparison between the adjusted $R^2$’s (percentage of total variability in the outcome that is accounted for by the entire set of predictors) across type-1 and corresponding type-2 models: type-2 models provide approximately 25% of the explanation for the overall Epi-QoL score, compared with 0.4, 1, and 2% for models with age at onset, age, and duration of epilepsy alone, respectively.”

We also modified the previous sentence in the following way: “For both continuous and categorical age-related models, seizure frequency in the preceding 12 months, psychiatric disturbances, number of medications, geographic area of the study center, gender, intellectual functions, and education explain most of the overall Epi-QoL score, and are, therefore, more relevant determinants of HRQOL (measured according to the overall Epi-QoL score) than age-related factors.”

Discussion.

Q26. Replace “is devoted to elucidate the role” with “is devoted to the elucidation of the role”.
A26. We modified the text as suggested.

Page 8.

Q27. Replace “revealed to” with “were found”.
A27. We modified the text as suggested.

Q28. The authors state: “Ageing is associated with a decline in learning and memory performances”. There is a tremendous spread of variability in older people with some maintaining good function and others not. The authors could overcome this by referring to a “mean” decline. In particular, in the discussion, the authors offer no reason why people with epilepsy should decline more with age than the general population and, it could be argued, that they have no reason for doing so since there is no control group, a fact that greatly weakens the scientific value of the study. The second half of this sentence is a gross over-generalisation: “although this decline was apparently similar for epilepsy patients and healthy subjects, the former group reached poor performance levels much earlier than the latter, as epilepsy patients fail to build up adequate learning and memory performance during childhood and adolescence”. Again, there is an enormous spread of variability; some childhood and adolescent syndromes are associated with gross cognitive deterioration whereas others are apparently not associated with such deterioration at all.
A28. We modified the first sentence referring to a “mean decline”. We modified the second sentence substituting “fail” with “may fail”.

Page 9.

Q29. Line 3. Replace “duration” with “found that duration”.
Replace “ot” with “to”.
Replace “on adults operated for“ with “adults who had operations for”.
Replace “On the contrary” with “On the contrary, the paper by” and delete the word “paper “after “co-authors”.
Replace “and other to” with “to other”.
Replace “in the long run” with “in the long term”.
A29. We modified the text as suggested.
Q30. Again, at this point, the paper lacks any discussion of the deterioration that is associated with some early-onset childhood epilepsy syndromes.
A30. We believe that the evidence on age at onset emerging from our analysis is only limited and we preferred not to discuss this topic in detail in the Discussion for this reason.

Page 10.
Q31. Where the authors have written “resulted to be relevant” do they mean “were found to be relevant”?
A31. Yes, we do, and we modified the text as suggested.

Q32. Where the authors have stated that early age of onset was associated with a lower quality of life, on this page they quote other work that apparently found “age at onset emerged as negatively associated with HRQOL”, implying that the higher the age of onset, the lower the quality of life. The last half of this sentence, before “Conclusions” might involve some discussion of this matter but this is not clear from what the authors have written.
A32. The mentioned work which is in apparent contradiction with the main analysis concerns a subgroup analysis on medication-resistant patients and the main point that emerged is that the results are nonsignificant. We believe that only statistically significant results are worthy to be discussed.

Q33. In the conclusions the authors have stated: “The main conclusions that age, age at onset and duration of epilepsy have only a limited role in determining the overall HRQOL”. With the confusion of different models apparently yielding different results, it would be helpful if the authors could reach some final conclusion that would be meaningful to workers in the field. At present the reader is left wondering what the degree, significance and importance of any correlations might be.
A33. We modified the conclusion in the following way: “The main conclusion is that age, age at onset, and duration of epilepsy have only a limited role in determining the overall HRQOL: age and duration of epilepsy show a generally significant negative effect (with the effect of duration being stronger and more consistent across models than the one of age), and age at onset a positive effect of limited significance, when considered alone or with another age-related factor, but, when other correlates of the overall Epi-QoL score are included in the regression models, any significant effect of the age-related factors disappears.”. 
Reviewer: Henry Potts
Reviewer's report:
I have focused on the statistical aspects of this paper. I don't know the epilepsy field well, so I can't comment on the originality of the work. (However, the review system requires me to select options for Level of interest and Quality of written English.)
As far as the statistics is involved, I have no complaints. The analyses performed are appropriate and described well. The attention to detail around the regression analyses is commendable. While I understand Dr Besag's concerns in his review about comprehensibility to a general readership, with my statistical hat on, I found the statistical detail useful and important.
In fact, I prefer the original wording around the authors' use of a Box-Cox transformation to the abbreviated version in the current draft.
We definitely agree with you, but we left the one of the current draft in consideration of Dr. Besag point of view.

I suggest two discretionary revisions.
1. The use of three decimal places for the results seems unnecessary. Two would suffice.
We left the three decimal places for p-values only and we removed them in all the other cases.
2. I would like to see more detail on the analyses adjusting for other confounders. Details are given for the analyses using categorical age-related variables, but I would like to see the analyses using the continuous age-related variables.
In the following, I reported the R software output for type-1 models.
Remember that: eta_new=age, durata=duration, eta_esordio=age at onset.
> m_eta<-lm(somma~ eta_new, data=pulito)
> summary(m_eta)

Call:
lm(formula = somma ~ eta_new, data = pulito)

Residuals:
     Min      1Q  Median      3Q     Max
-144.679  -29.046    8.837   34.837   72.264

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)     217.10541    4.46854  48.585   <2e-16 ***
eta_new         -0.24174    0.10771  -2.244   0.0251 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 . ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.55 on 813 degrees of freedom
Multiple R-squared: 0.006153,  Adjusted R-squared: 0.004931
F-statistic: 5.034 on 1 and 813 DF,  p-value: 0.02513

> m_dur<-lm(somma~ durata, data=pulito)
> summary(m_dur)

Call:
lm(formula = somma ~ durata, data = pulito)

Residuals:
   Min   1Q Median   3Q   Max
-142.399 -27.090  8.145 33.954 72.720

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  216.3409    2.6569  81.426  < 2e-16 ***
durata       -0.4559    0.1149  -3.968 7.89e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.27 on 813 degrees of freedom
Multiple R-squared: 0.019, Adjusted R-squared: 0.01779
F-statistic: 15.75 on 1 and 813 DF, p-value: 7.888e-05

> m_esor<-lm(somma ~ eta_esordio, data=pulito)
> summary(m_esor)

Call:
  lm(formula = somma ~ eta_esordio, data = pulito)

Residuals:
   Min   1Q Median   3Q   Max
-144.025 -28.906   8.922  34.239  68.239

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  205.05008    2.51432  81.553   <2e-16 ***
eta_esordio   0.13163    0.09983   1.319    0.188
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.64 on 813 degrees of freedom
Multiple R-squared: 0.002134, Adjusted R-squared: 0.0009065
F-statistic: 1.739 on 1 and 813 DF, p-value: 0.1877

> m_eta_dur<-lm(somma ~ eta_new + durata, data=pulito)
> summary(m_eta_dur)

Call:
  lm(formula = somma ~ eta_new + durata, data = pulito)

Residuals:
   Min   1Q Median   3Q   Max
-142.591 -28.055  7.981 33.997 72.907

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  219.34848    4.49013  48.851  < 2e-16 ***
eta_new  -0.09595   0.11547  -0.831 0.406236
durata   -0.41731   0.12397  -3.366 0.000798 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.28 on 812 degrees of freedom
Multiple R-squared: 0.01983,  Adjusted R-squared: 0.01742
F-statistic: 8.215 on 2 and 812 DF,  p-value: 0.0002937

> m_eta_esor<-lm(somma~eta_new+eta_esordio, data=pulito)
> summary(m_eta_esor)

Call:
  lm(formula = somma ~ eta_new + eta_esordio, data = pulito)

Residuals:
   Min     1Q   Median     3Q    Max
-142.591  -28.055    7.981   33.997   72.907

Coefficients:
    Estimate Std. Error t value Pr(>|t|)
(Intercept)  219.3485    4.49013  48.851  < 2e-16 ***
eta_new      -0.5133     0.13404  -3.829 0.000139 ***
eta_esordio   0.4173     0.12403   3.366 0.000798 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.28 on 812 degrees of freedom
Multiple R-squared: 0.01983,  Adjusted R-squared: 0.01742
F-statistic: 8.215 on 2 and 812 DF,  p-value: 0.0002937

> m_dur_esor<-lm(somma~durata+eta_esordio, data=pulito)
> summary(m_dur_esor)

Call:
  lm(formula = somma ~ durata + eta_esordio, data = pulito)

Residuals:
   Min     1Q   Median     3Q    Max
-142.591  -28.055    7.981   33.997   72.907

Coefficients:
    Estimate Std. Error t value Pr(>|t|)
(Intercept)  219.34848    4.49013  48.851  < 2e-16 ***
durata      -0.51326     0.13404  -3.829 0.000139 ***
eta_esordio -0.09595     0.11547  -0.831 0.406236
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 43.28 on 812 degrees of freedom
Multiple R-squared: 0.01983,  Adjusted R-squared: 0.01742
In the following, I reported the R software output for type-2 models. Confounders appeared in the same order as in tables 2 and 3.

```r
> m_eta_fin <- update(m_eta, ~ . + n_cris_12m_new + dis_psi_new + Sesso + n_terapie_new + liv_int_new + scol + rip_geo + eziol_new, data = pulito)
> summary(m_eta_fin)

Call:
  lm(formula = somma ~ eta_new + n_cris_12m_new + dis_psi_new + Sesso + n_terapie_new + liv_int_new + scol + rip_geo + eziol_new, data = pulito)

Residuals:
  Min       1Q   Median       3Q      Max
-121.178  -23.505    5.568   26.565   93.531

Coefficients:
                        Estimate Std. Error t value   Pr(>|t|)
(Intercept)              217.80112   6.71257  32.447  < 2e-16 ***
eta_new                   -0.07924   0.10331  -0.767 0.443291
n_cris_12m_new1-5        -18.48486   3.47428  -5.320 1.35e-07 ***
n_cris_12m_new6-20       -23.18865   4.47159  -5.186 2.73e-07 ***
n_cris_12m_new>20        -32.67120   4.01928  -8.129 1.65e-15 ***
dis_psi_newYes           -20.12827   4.03134  -4.993 7.30e-07 ***
SessoM                    8.27852    2.70502   3.060 0.002284 **
n_terapie_newMonotherapy  9.78276    3.02374   3.235 0.001265 **
n_terapie_newNo therapy  19.84463   10.87670   1.825 0.068449 .
liv_int_newCompromised   -12.31233    4.09953  -3.003 0.002754 **
  scol9-13 years           6.71226    3.14319   2.135 0.033024 *
  scol>=14 years           9.94300    4.15617   2.392 0.016971 *
  rip_geoCentral Italy    -2.37152    3.28329  -0.722 0.470322
  rip_geoSouthern Italy   -13.06822    3.46216  -3.775 0.000172 ***
eziol_newCryptogenic     4.06024    3.08253   1.317 0.188157
eziol_newIdiopathic      3.29496    3.83994   0.858 0.391108
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 37.77 on 799 degrees of freedom
Multiple R-squared: 0.2653,  Adjusted R-squared: 0.2515
F-statistic: 19.24 on 15 and 799 DF,  p-value: < 2.2e-16

> m_dur_fin <- update(m_dur, ~ . + n_cris_12m_new + dis_psi_new + Sesso + n_terapie_new + liv_int_new + scol + rip_geo + eziol_new, data = pulito)
> summary(m_dur_fin)

Call:
  lm(formula = somma ~ durata + n_cris_12m_new + dis_psi_new + Sesso + n_terapie_new + liv_int_new + scol + rip_geo + eziol_new, data = pulito)

Residual standard error: 37.77 on 799 degrees of freedom
Multiple R-squared: 0.2653,  Adjusted R-squared: 0.2515
F-statistic: 19.24 on 15 and 799 DF,  p-value: < 2.2e-16
```
Sesso + n_terapie_new + liv_int_new + scol + rip_geo + eziol_new, data = pulito)

Residuals:
    Min   1Q Median   3Q   Max
-122.216 -23.928   5.955  26.392  94.205

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 215.77389    5.60134  38.522  < 2e-16 ***  
durata      -0.06006    0.11147  -0.539 0.590196  
n_cris_12m_new1-5  -18.84079    3.48320  -5.409 8.38e-08 ***  
n_cris_12m_new6-20  -23.48093    4.47572  -5.246 1.99e-07 ***  
n_cris_12m_new>20  -32.63608    4.02001  -8.118 1.78e-15 ***  
dis_psi_newYes    -20.38137    4.01047  -5.082 4.65e-07 ***  
SessoM          7.99455    2.70562   2.955 0.003221 **  
n_terapie_newMonotherapy  9.40560 3.13705  2.998 0.002800 **  
n_terapie_newNo therapy  19.40671 10.95105  1.772 0.076754 .  
liv_int_newCompromised   -11.82156    4.12526  -2.866 0.004271 **  
scol9-13 years        7.11779    3.06243   2.324 0.020363 *  
scol>=14 years        10.36891    4.08544   2.538 0.011337 *  
rip_geoCentral Italy -2.43885   3.29900  -0.739 0.459961  
rip_geoSouthern Italy 12.74696   3.41320   3.735 0.000201 ***  
eziol_newCryptogenic  4.21104   3.07622   1.369 0.171415  
eziol_newIdiopathic  3.91875   3.77638   1.038 0.299723  

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 .’ 0.1 ‘ ’ 1

Residual standard error: 37.78 on 799 degrees of freedom  
Multiple R-squared: 0.2651, Adjusted R-squared: 0.2513  
F-statistic: 19.21 on 15 and 799 DF, p-value: < 2.2e-16

> ### m_esor_c:  
> m_esor_fin<-update(m_esor, . ~ .+n_cris_12m_new+dis_psi_new+Sesso+n_terapie_new+liv_int_new+scol+rip_geo+eziol_new, data=pulito)  
> summary(m_esor_fin)

Call:  
    lm(formula = somma ~ eta_esordio + n_cris_12m_new + dis_psi_new + Sesso + n_terapie_new + liv_int_new + scol + rip_geo + eziol_new, data = pulito)

Residuals:
    Min   1Q Median   3Q   Max
-122.587 -23.671   6.131  26.343  94.324

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 214.46936    4.90730  43.704 < 2e-16 ***  
eta_esordio -0.02339    0.09507  -0.246 0.805734
n_cris_12m_new1-5  -18.53962  3.50303  -5.292 1.56e-07 ***
n_cris_12m_new6-20  -23.24041  4.48700  -5.180 2.82e-07 ***
n_cris_12m_new>20  -32.66155  4.02079  -8.123 1.72e-15 ***
dis_psi_newYes  -20.40841  4.01628  -5.081 4.67e-07 ***
SessoM  8.20701  2.72464  3.012 0.002676 **
 n_terapie_newMonotherapy 10.01169  3.08725  3.243 0.001232 **
n_terapie_newNo therapy 20.30404  10.90383  1.862 0.062956 .
liv_int_newCompromised -12.27727  4.15107  -2.958 0.003192 **
scol9-13 years  7.25035  3.05886  2.370 0.018011 *
scol>=14 years  10.58124  4.06459  2.603 0.009405 **
rip_geoCentral Italy  -2.21029  3.28404  -0.673 0.501116
rip_geoSouthern Italy  -12.56038  3.39231  -3.703 0.000228 ***
eziol_newCryptogenic  4.17471  3.08178  1.355 0.175914
 eziol_newIdiopathic  3.65110  3.85355  0.947 0.343689
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 37.79 on 799 degrees of freedom
Multiple R-squared: 0.2648,  Adjusted R-squared: 0.251
F-statistic: 19.19 on 15 and 799 DF,  p-value: < 2.2e-16

> ### m_eta_dur_c:
> m_eta_dur_fin<update(m_eta_dur, . ~

> .+n_cris_12m_new+dis_psi_new+Sesso+n_terapie_new+liv_int_new+scol+rip_geo+eziol_new, data=pulito)
> summary(m_eta_dur_fin)

Call:
  lm(formula = somma ~ eta_new + durata + n_cris_12m_new + dis_psi_new +
  Sesso + n_terapie_new + liv_int_new + scol + rip_geo + eziol_new,
  data = pulito)

Residuals:
  Min 1Q Median 3Q Max
-121.071 -23.701  5.736  26.484  93.508

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  218.22835   6.89061  31.670  < 2e-16 ***
eta_new     -0.06797    0.11106  -0.612  0.540666
durata      -0.03325    0.11980  -0.278  0.781445
n_cris_12m_new1-5  -18.60921    3.50504  -5.309 1.43e-07 ***
n_cris_12m_new6-20  -23.28820    4.48853  -5.188 2.69e-07 ***
n_cris_12m_new>20  -32.66053    4.02178  -8.121 1.75e-15 ***
dis_psi_newYes  -20.12600    4.03368  -4.989 7.43e-07 ***
SessoM         8.18978    2.72540   3.005 0.002676 **
n_terapie_newMonotherapy  9.54310    3.14631  3.033 0.002499 **
n_terapie_newNo therapy 20.30404    10.90383  1.949 0.052556 .
liv_int_newCompromised  -12.27727    4.15107  -2.958 0.003192 **
scol9-13 years   6.67508    3.14786   2.121 0.034271 *
scol>=14 years    9.87646    4.16548   2.371 0.017975 *
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 . ‘1

Residual standard error: 37.79 on 798 degrees of freedom
Multiple R-squared: 0.2654,  Adjusted R-squared: 0.2507
F-statistic: 18.02 on 16 and 798 DF,  p-value: < 2.2e-16

> m_eta_esor_fin<-update(m_eta_esor, . ~
+.n_cris_12m_new+dis_psi_new+Sesso+n_terapie_new+liv_int_new+scol+rip_geo+eziol_new, data=pulito)
> summary(m_eta_esor_fin)

Call:
  lm(formula = somma ~ eta_new + eta_esordio + n_cris_12m_new +
      dis_psi_new + Sesso + n_terapie_new + liv_int_new + scol +
      rip_geo + eziol_new, data = pulito)

Residuals:

     Min      1Q  Median      3Q     Max
-121.071 -23.701    5.736   26.484   93.508

Coefficients:
             Estimate Std. Error  t value Pr(>|t|)
(Intercept)  218.22835   6.89061  31.670  < 2e-16 ***
eta_new      -0.10122   0.13023  -0.777  0.437217
eta_esordio   0.03325   0.11980   0.278  0.781445
n_cris_12m_new1-5  -18.60921  3.50504  -5.309  1.43e-07 ***
n_cris_12m_new6-20  -23.28820  4.48853  -5.188  2.69e-07 ***
n_cris_12m_new>20  -32.66053  4.02178  -8.121  1.75e-15 ***
dis_psi_newYes  -20.12600  4.03368  -4.989  7.43e-07 ***
SessoM        8.18978   2.72540   3.005  0.002739 **
n_terapie_newMonotherapy  9.54310   3.14631  3.033  0.002499 **
n_terapie_newNo therapy  19.49357  10.95626  1.779  0.075585
liv_int_newCompromised  -12.12567  4.15667  -2.917  0.003632 **
scol9-13 years  6.67508   3.14786   2.121  0.034271 *
scol14 years   9.87646   4.16548   2.371  0.017975 *
rip_geoCentral Italy  -2.45955  3.30047  -0.745  0.456362
rip_geoSouthern Italy  -13.12834  3.47093  -3.782  0.000167 ***
eziol_newCryptogenic   4.07861   3.08502   1.322  0.186525
eziol_newIdiopathic    3.41506   3.86645   0.883  0.377364

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 37.79 on 798 degrees of freedom
Multiple R-squared: 0.2654,  Adjusted R-squared: 0.2507
F-statistic: 18.02 on 16 and 798 DF,  p-value: < 2.2e-16
```r
> m_dur_esor_c:
> m_dur_esor_fin<-update(m_dur_esor, . ~ .+n_cris_12m_new+dis_psi_new+Sesso+n_terapie_new+liv_int_new+scol+rip_geo+eziol_new, data=pulito)
> summary(m_dur_esor_fin)

Call:
  lm(formula = somma ~ durata + eta_esordio + n_cris_12m_new +
      dis_psi_new + Sesso + n_terapie_new + liv_int_new + scol +
      rip_geo + eziol_new, data = pulito)

Residuals:
   Min     1Q   Median     3Q    Max
-121.071 -23.701    5.736   26.484   93.508

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   218.22835    6.89061  31.670  < 2e-16 ***
durata         -0.10122    0.13023  -0.777 0.437217
eta_esordio    -0.06797    0.11106  -0.612 0.540666
n_cris_12m_new1-5  -18.60921    3.50504  -5.309 1.43e-07 ***
n_cris_12m_new6-20   -23.28820    4.48853  -5.188 2.69e-07 ***
n_cris_12m_new>20    -32.66053    4.02178  -8.121 1.75e-15 ***
dis_psi_newYes    -20.12600    4.03368  -4.989 7.43e-07 ***
SessoM            8.18978    2.72540   3.005 0.002739 **
n_terapie_newMonotherapy  9.54310    3.14631   3.033 0.002499 **
n_terapie_newNo therapy 19.49357   10.95626   1.779 0.075585 .
liv_int_newCompromised -12.12567    4.15667  -2.917 0.003632 **
scol9-13 years     6.67508    3.14786   2.121 0.034271 *
scol>=14 years      9.87646    4.16548   2.371 0.017975 *
rip_geoCentral Italy -2.45955    3.30047  -0.745 0.456362
rip_geoSouthern Italy -13.12834   3.47093  -3.782 0.000167 ***
eziol_newCryptogenic  4.07861    3.08502   1.322 0.186525
eziol_newIdiopathic  3.41506    3.86645   0.883 0.377364
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 37.79 on 798 degrees of freedom
Multiple R-squared: 0.2654,  Adjusted R-squared: 0.2507
F-statistic: 18.02 on 16 and 798 DF,  p-value: < 2.2e-16
```