Author's response to reviews

Title: Patients in long-term maintenance therapy for drug use in Italy: analysis of some parameters of social integration and serological status for infectious diseases in a cohort of 1091 patients.

Authors:

Gianluca Quaglio (gianluca.quaglio@azosp.vr.it)
Fabio Lugoboni (fabio.lugoboni@azosp.vr.it)
Cristian Pattaro (cristian@biometria.univr.it)
Gics Investigators (medicina.dipendenze@azosp.vr.it)
Linda Montanari (Linda.Montanari@emcdda.org)
Alessandro Lechi (alessandro.lechi@univr.it)
Paolo Mezzelani (paolo.mezzelani@univr.it)
Don CD Jarlais (dcdesjarla@aol.com)

Version: 4 Date: 19 April 2006

Author's response to reviews: see over
Dear BioMed Central Editorial Team,

Thank you for your answer. I am providing you with a point-by-point response to the reviewer and the statistician, and the manuscript revised. New phrases and sentences removed from the original text are in **bold** type.

We would like to comment on observations (**in italics**) as follows:

**COMMENTS ON REVIEWER OBSERVATIONS**

1) Previous comment 3: “The description of table 4 in the text is flawed. For instance, older subjects were not only at increased risk of hepatitis and TPHA but also HIV. Employment status is also incorrectly described. Furthermore it is not clear whether the authors included all variables in a logistic regression (LR) for each infectious disease or whether they are all different LR (adjusted by age –continuous variable?- and sex).”

Description of table 4 is now correct, although the way the table is constructed is misleading. The legend under table 4 doesn’t clarify its content. Figures in cells for each column are not either univariate, nor all of them from a multivariate model, whereas according to what the authors say (in the accompanying letter) there are two multivariate models in each column. This is absolutely estrange as some of the cells are repeated in each model (i.e.; age, sex), and one can not know to which one do the coefficients belong. Furthermore I do not agree with authors when they say “The reason is that it is not clear if and how employment/education could confound or be confused by partner variables for serology”. A wide body of knowledge exist around inequalities in health. Education and employment are some of the variables used to ascertain social position and such measure inequalities in health. Thus a multivariate model with all variables would be absolutely coherent.

Table 4 and related statistical analyses have now been modified: four logistic models have been fitted, one for each serological trait. All the models were multivariate, meaning that all the variables reported in the table were included in the models together, plus a factor variable related to the recruitment centers. When including all these factors together, a lack of power could affect the estimation: for these reason we compared models adjusted by recruitment unit, with model without this adjustment and results looked to be consistent. The model we used is now clearly reported in the Statistical analyses section. The section has been revised and modelisation is now better explained as follows:

To ensure the consistency of the comparison with the general population, of a slightly different age, the direct standardisation method was used. To assess the association of social integration variables with the serological variables the following logistic regression models was fitted (one for each serological variable):  

$$ \text{logit}(\theta) = \alpha + \beta_{1,1.15} \text{(recruitment unit)} + \beta_2 \text{(age)} + \beta_3 \text{(sex)} + \beta_{4,1.3} \text{(education)} + \beta_{5,1.2} \text{(employment status)} + \beta_{6,1.3} \text{(living status)} + \beta_{7,1.2} \text{(partner status)} + \beta_8 \text{(partner’s use of heroin)} + \varepsilon$$

where: $\theta$ was the probability of HIV, HBV, HCV and TPHA, respectively. All the independent variables were included as factors (age was dichotomised into $<35$ and $\geq 35$ years; for education, the degree level was omitted due to the too small number of subjects); $\varepsilon$ was a random error following a binomial distribution.

As a consequence of the changes of table 4, in Results section, the subsection Social characteristics and serological status has been modified as follow: **As shown in Table 4, older subjects had an increased risk of all the diseases (OR ranged within 2.55 and 6.25). Males had a lower risk of TPHA (OR=0.28; 95% CI: 0.08,0.98) than females. Higher education was associated with a lower prevalence of HIV: OR=0.34 (95% CI: 0.14,0.84) for secondary school with respect to primary education. The status of unemployment was associated with an increased prevalence of HIV (OR=2.05, 95% CI: 1.26-3.34) and HCV (OR=1.72, 95% CI:
1.02-2.94). Subjects living with a partner had a higher risk of HIV and HCV than subjects living with parents: OR=2.51 (95% CI: 1.13, 5.56) and OR=1.75 (95% CI: 1.05-2.94), respectively. Subjects with stable relationships looked to have a lower risk of HIV (OR=0.28, 95% CI: 0.10-0.75 for married vs no steady partner subjects) and of HBV (OR=0.55, 95% CI: 0.32-0.96 for steady vs unsteady partner subjects).

2) Previous comment 4: May be I misunderstand the sentence “Patients with good educational level seem more motivated in relation to changing their behaviour” (bottom page 9), but I don’t agree with the authors’ response. I think the sentence is not accurate as there is no information about the motivation of the patients in this respect. The only thing observed was that patients with higher education had a shorter length of use, after being an unknown length of time into treatment. Probably what could be inferred is that they “seem motivated earlier”.

Following the suggestion of the referee, the phrase, now in page 10, has been modified as follow: **Patients with good educational levels seem motivated earlier in relation to changing their behaviour.**

3) Previous comment 5: The change the authors propose in their letter has not been corrected in the manuscript. Thank you. The error, in the subsection Living and partner status, page 11, has been correct.

**COMMENTS ON STATISTICIAN OBSERVATIONS**

1) I am in agreement with the reviewer that the presentation of Table 4 appears to be misleading given the explanation in the Author’s response to the reviews. Table 4 appears to show 4 models - one for each of the seriological conditions. Each model appears to have all of the variables in the first column included. This is at odds with the explanation given in the Author’s response. Table 4 should show a separate column for each model. All of the variables included in each model should be presented. If Table 4 is showing only 4 models - for the 4 seriological conditions - then, for example, it is not clear how partner's drug use was incorporated into each model since only those people with a partner could answer that question. More explanation is needed. The legend for Table 4 also does not make sense. What does it mean for the “results to be consistent” or for variables to be “mutually adjusted”. More explanation in the methods section is needed.

As described above, table 4 have now been modified. All the models were multivariate. When including all these factors together, a lack of power could affect the estimation: for these reason we compared models adjusted by recruitment unit, with model without this adjustment and results looked to be consistent. An unadjusted version of table 4 (Table 4bis) is included below for the reviewers.

The expression “mutually adjusted” was omitted. Anyway, it is a generally accepted expression in the epidemiological slang: it means that different factors were included together in the model, so the adjustment is, in fact, one factor with respect the others.

Table 4bis. This is the same of table 4, but the variable “recruitment unit” which consisted of 16 levels, was omitted.
### Secondary 0.30 (0.12, 0.72) 0.52 (0.25, 1.08) 0.72 (0.36, 1.43) 0.29 (0.02, 3.68)

#### Employment

<table>
<thead>
<tr>
<th></th>
<th>Employment</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employed</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Unemployed</td>
<td>2.21</td>
<td>1.70</td>
<td>1.88</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>3.85</td>
<td>1.09</td>
<td>0.58</td>
</tr>
</tbody>
</table>

#### Living status

<table>
<thead>
<tr>
<th></th>
<th>Parents</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Partner</td>
<td>2.74</td>
<td>1.61</td>
<td>1.97</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Alone</td>
<td>0.67</td>
<td>1.83</td>
<td>1.25</td>
<td>n.e.</td>
</tr>
<tr>
<td></td>
<td>Friends</td>
<td>1.05</td>
<td>0.94</td>
<td>1.38</td>
<td>0.87</td>
</tr>
</tbody>
</table>

#### Partner

<table>
<thead>
<tr>
<th></th>
<th>No steady</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Steady</td>
<td>0.58</td>
<td>0.67</td>
<td>0.62</td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>Married</td>
<td>0.24</td>
<td>0.62</td>
<td>0.49</td>
<td>0.47</td>
</tr>
</tbody>
</table>

#### Partner’s drug use

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>1.50</td>
<td>1.19</td>
<td>1.31</td>
<td>0.21</td>
</tr>
</tbody>
</table>

2) Other notes:
For all of the regression analyses done in the paper goodness of fit statistics should be reported. Other model diagnostics should be reported for the linear regression in particular. It is not clear that normality is a reasonable assumption for duration of drug use.

The distribution of duration of drug use is reported in the following Figure: it is not normal distributed (p-value of the Shapiro test for normality was <0.05), but it is quite symmetrical (median=14.00, mean=14.54 years). We unsuccessfully tried many transformations to make the data normally distributed. The Bartlett’s test for homogeneity of variances across groups, never rejected the hypothesis of same variances, but the employment factor.

Thus, we agree that caution is needed when reading the confidence intervals, but this was the best we could do. Goodness of fit plots for the model “duration ~ recruitment unit + age + sex” are reported below. We observed a lack of normality in the residuals, but also a certain symmetry on the estimated residuals vs the fitted data, thus suggest that the estimates should not be too much biased.
To support the findings, a column with the p-values of the LRS statistics for the inclusion of the factors into the model with recruitment centre, age and sex, was added to the Table 2.

As consequence of the modification of Table 2, in Results section, subsection Drug use characteristics, has been modified as follow: …… A relevant amount of the total variability was explained by age, sex and recruitment centre (adjusted-R²=0.61). Additional factors explaining a significant amount of variation were education (LRS p-value<0.0001) and employment status (LRS p-value=0.0340). In fact, after adjusting by age, sex and recruitment centre, patients with a vocational school education had used drugs, on average, 1.2 years less than those with a primary school (p-value=0.05); subjects with a secondary school education used drugs on average for 2.2 years less than patients with only primary school education (p<0.001). The unemployed had used drugs for an average of 1 year longer than those in employment (p=0.012).
3) Some of the percentages in Table 1 are not correct. The authors should also indicate, with a symbol say, which p-values came from Fisher’s exact tests; simply saying “used where cell numbers were too small” is too ambiguous. We have found 1 erroneous percentage. We better explained the table and fixed other minor errors. A double * has been added to recognised the p-values from the Fisher’s exact test.

4) Another question since I feel it is very important in such studies. Was ethics committee approval given for this study? I did not notice this in the paper and I feel it is a very important consideration given the sensitive nature of the data.

In Methods, in the section Population, the following phrase has been added: The study has been approved by the local Ethical Committee.

I hope these corrections will answer the referees objections and satisfy the Editorial Board.

Looking forward to hearing from you.

Yours sincerely,

Gianluca Quaglio, MD