INTRODUCTION

Hepatitis C virus (HCV) is a leading cause of end-stage liver disease throughout the world. The World Health Organization estimates that in Europe about 9 million people live with HCV and that the virus produces about 86,000 deaths per year, figures which more than double those estimated for hepatitis B (WHO 2013). Estimates for North America are quite similar and recently published data indicates that about 5.2 million people are infected with HCV in the USA (Chak et al., 2011). Sustained virological response (SVR) is proved to reduce overall and liver-related mortality in patients with HCV.

In particular, a recent systematic review of 18 prospective studies found that, in comparison to viraemic patients, patients who achieve SVR have an adjusted hazard ratio between 0.07 and 0.39 and between 0.04 and 0.27 for all-cause and liver-related mortality, respectively (Chou et al., 2013).
Since the early 2000s, dual therapies with ribavirin and pegylated alpha-interferons (either peg-IFN alpha-2a or peg-IFN alpha-2b) have become the standard of care (SoC) for patients chronically infected with HCV. However, the success of achieving SVR under these treatments significantly depends on several factors, such as previous response to interferon and viral genotype (Ghany et al., 2011; EASLD, 2012). In fact, while dual therapy can produce SVR in 80% of naïve patients infected with genotypes 2 and 3, the proportion of SVR in patients infected with genotype 1 is normally less than 50%. Figures for patients who failed previous treatment might be even more disappointing, being about 5% in null-responders (Ghany et al., 2009).

Treatment for HCV infection is rapidly evolving and several amazing new therapies will soon offer hope for difficult-to-treat patients. In particular, on the grounds of several successful clinical trials, two direct-acting antiviral (DAA) compounds, telaprevir and boceprevir, in association with peg-IFN alpha and ribavirin (i.e.: triple therapy) have recently been approved for treatment of HCV genotype 1 (Ghany et al., 2011; EASLD, 2012). The aim of this study was to measure the relative effects (improved SVR and increased occurrence of adverse events) of triple therapy in comparison with SoC through an extended frequentist meta-analytic approach. In addition, we considered non-standard meta-analysis techniques, such as generalized least square for trend estimation, to supplement direct comparisons of standard meta-analyses. The review is reported according to the "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement" (Liberati et al., 2009).

MATERIALS AND METHODS

Definitions

The following definitions are used in the present study. SVR: undetectable HCV-RNA, 24 weeks after completion of therapy; serious adverse event (SAE), severe infections and severe anaemia are defined according to the medical dictionary for regulatory activities as implemented in each specific study; treatment-naïve: patients who never received anti-HCV treatment; treatment-experienced: patients who received anti-HCV treatment. Null-responders: experienced patients who never had a reduction of less than 2 log in HCV RNA while in therapy; SoC: 48 weeks of peg-IFN plus ribavirin. Response-guided therapy (RGT): any reduction of treatment undertaken on the grounds of the patient’s response.

Eligibility criteria

We have included only randomized controlled trials (RCT) published in peer-reviewed journals and reports in a recognized trial registry until 15 April 2013 in English, Spanish, French, German and Italian. The population study had to include only adult patients (aged 18 or more) who were infected with genotype 1 HCV. Studies must have compared a SoC with any combination of peg-IFN alpha and ribavirin plus either boceprevir or telaprevir. Ribavirin dose reduction during treatment was allowed. We have excluded interim analyses, post-hoc analyses of larger RCT and any study which enrolled HIV or HBV co-infected subjects. In addition, intervention arms which did not include ribavirin are described in the systematic review but were not included in the inferential analyses.

Information source and search strategy

An electronic search was performed in PubMed, the Cochrane Controlled Trials Register and the NIH National Clinical Trial Registry (NCT). To provide the search strategy with maximum sensitivity we kept the searching string as simple as possible using only keys words without filters. The following string was used in all databases:

```
• "telaprevir" OR "incivek" OR "incivo" OR "VX950" OR "VX?950*" OR "MP?424" OR "boceprevir" OR "victrelis" OR "Sch?503034*" OR "Sch503034".
```

Particular attention was paid to retrieving the original protocol of each study and re-defining original study datasets throughout data mining.

Study selection and data collection

Study selection and data collection were carried out by two independent reviewers (SL and AM). A single paper was either included or dis-
carded if both reviewers independently reached the same decision. In case of discordant decisions, papers were assessed again by both reviewers for a consensus decision.

Data collected include: the number of patients randomized/treated; number of patients who achieved SVR; number of patients who experienced any SAE; number of patients who experienced severe anaemia; number of patients who experienced severe infections; dosage and time of exposure to each DAA; participants’ mean age; participants’ mean body mass index (BMI); participants’ mean HCV viral load; number of male participants; number of black participants; number of participants infected with HCV genotype 1a; number of patients with cirrhosis, previous treatment response to previous treatment; methodological issues relevant for the assessment of risk of bias.

Assessment of risk of bias
Potential risk of bias within the study was evaluated according to 5 different items; i.e.: sequence generation, allocation concealment, blinding, attrition and early termination. Sequence generation was adequate if either a random number table, stratification, a computer to generate a random sequence or minimization was used. Allocation concealment was adequate if either central allocation, sequentially numbered drug containers of identical appearance or sequentially numbered opaque sealed envelopes were used. Blinding was adequate if an appropriate placebo was used. Attrition was negligible if all the following criteria were met: a) randomized but untreated patients and overall lost to follow-up was less than 20%; b) lost to follow-up was balanced between study arm (i.e. chi-squared p-value for balance >0.05); c) there were no major discrepancies between data as reported in the published paper and trial registry report. Early termination was always considered a potential source of bias (Montori et al., 2005; Bassler 2010). A study was considered at low risk of bias if all the criteria were met, otherwise it was considered at a potentially high risk of bias.

The across-studies risk of bias was assessed by the analyses of small study effect using the Harbord test.

Assessment of the quality of a body of evidence
The quality of the evidence was assessed according to the “Grades of Recommendation, Assessment, Development and Evaluation Working Group approach (GRADE)” (Atkins et al., 2004). In brief, the quality of evidence was rated according to four levels of quality: high, moderate, low and very low. Quality level was assigned by a three-step procedure. Firstly, all outcomes are considered to be supported by high quality evidence, as we included only RCTs. Secondly, the level of quality was downgraded according to five criteria:
1) within-study risk of bias, present if less than one study supporting the evidence was at low risk of bias;
2) imprecision of results, present if a meta-analysis did not meet the criteria for optimal information size (OIS) (Guyatt; et al., 2011a);
3) indirectness of evidence, present if indirect comparison was made;
4) across the study risk of bias, present if the Harbord test provided evidence of significant effect of small studies (Guyatt; et al., 2011b);
5) inconsistency of results, present if the meta-analysis had unexplained heterogeneity (I2>20%). Finally, quality of evidence for outcomes was upgraded according to effect size: a grade of +1 if the relative effect was between 2-5 or 0.5-0.2; a grade of + 2, if the relative effect was >5 or <0.2.

Additional analyses and statistical methods
The Risk ratio (RR) was used as a measure of effect. As most of the studies are multi-arm RCTs, we combined groups in order to create a single pair-wise comparison, as recommended by the Cochrane collaboration guidelines (Higgins et al., 2011). All arms in a single study which contained DAA are combined and compared to the SoC. Meta-analyses were carried out according to the inverse variance model and DerSimonian and Laird random effect estimates were always provided as terms of comparison. For all meta-analyses with I2 >20% we investigated the potential cause of heterogeneity according to population baseline characteristics. Subgroup analysis was used to investigate hetero-
geneity for categorical variables and the simple meta-regression model was used to explore potential causes of heterogeneity for continuous variables. The multiple meta-regression model was used to assess the simultaneous effect of multiple covariates. Generalized least square for trend estimation, as reported elsewhere (Greenland et al., 1992; Berilin et al., 1993), was used to assess dose response for SVR and SAE. Sensitivity analysis was undertaken to assess robustness of evidence. The STATA 11.1 (StataCorp Texas 77845 USA) package was used for the analysis and to generate plots.

RESULTS

Study selection
The results of the study selection are reported in Figure 1.

Systematic review
The systematic review includes 10 RCTs comprising 33 individual treatment arms which enrolled 5,312 participants, 5,186 of whom (97.63%) received at least one dose of treatment. Baseline population features varied significantly between studies with regard to several variables potentially associated with SVR, such as the proportion of black patients and patients with cirrhosis. Table 1 summarizes the patients’ baseline characteristics.

Six studies are RCTs comparing the SoC with telaprevir regimens, 2 of which include treatment-experienced patients and null-responders. Four studies are RCTs comparing the SoC with boceprevir regimens; 2 of these include patients who did not achieve SVR in a previous treatment but none of them include null-responders. Two out of the 23 intervention arms (78 and 111 patients respectively) consider ribavirin-free regimens; these arms are described in the systematic review but are not included in the meta-analyses.

Table 2 reports the overall summary of patient outcomes according to SVR, SAE, severe anaemia and severe infection. Triple therapy with DAA is almost always superior to SoC with regard to SVR.

The individual RRs for SVR in patients receiving triple therapy range between 1.29 and 3.98. Only 3 out of the 23 interventions arms do not provide statistically significant evidence that DAA regimens are superior to SoC; i.e.: the 2 ribavirin-free arms and one ultra-short triple therapy with telaprevir (12-week overall therapy).

Evidence for increased risk of adverse events due to triple therapy with DAA is contrasting. Individual RRs for SAEs, severe anaemia and severe infections range widely and are only occasionally significant. In addition, there is no evidence that avoidance of ribavirin is associated with a decreased risk of any SAE.

Within-study risk of bias
None of the studies are judged to be at low risk of bias.

Random sequence generation was adequate in all the studies. Patient allocation methods were adequate in 7 studies while 3 studies do not clearly report how patients were allocated. Five
studies were blinded, two were open label and 3 were only partially blinded. In particular, the ADVANCE study was considered as partially blinded, since both intervention arms were unmasked by the 24th week of treatment; the result of the study design which considers RGT for both intervention arms and fixed peg-IFN and ribavirin therapy for the control arm. Attrition was considered negligible in 5 studies. Two studies (Flamm 2012 and SPRINT-2) were at high risk of bias due to imbalanced loss to follow-up, and one study (RESPOND 2) was at unclear risk of bias since we could not assess attrition due to major discrepancies between data reported in the NCT registry and data published in the peer-reviewed paper. No study was terminated early.

**Results of the meta-analyses**

Figure 2 shows the forest plots with results of the meta-analyses for SVR (10 studies), SAE (9 studies), severe anaemia (9 studies) and severe infections (9 studies). The SPRINT-1 study was only included in SVR meta-analyses as patients with detectable HCV-RNA levels at week 24 of treatment were allowed to opt for boceprevir plus peg-IFN and ribavirin for a total treatment duration of 54 weeks. This made the data on adverse reaction endpoints unreliable.

Meta-analysis for SVR (Figure 2A) included 4,997 patients. SVR was achieved in 2,315 out of the 3,556 patients who received DAA, and in 503 out of the 1,441 patients who received SoC.

### TABLE 1 - Description of population baseline characteristics in the 10 RCTs included in the systematic review and meta-analyses.

<table>
<thead>
<tr>
<th>Study</th>
<th>Randomized Treated (N %)</th>
<th>Population</th>
<th>Black (N %)</th>
<th>Gen. 1a (N %)</th>
<th>Cirr. (N %)</th>
<th>Male (N %)</th>
<th>Viral load (mean log)</th>
<th>BMI (mean)</th>
<th>Age (mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advance</td>
<td>1095</td>
<td>1088</td>
<td>Naïve</td>
<td>94</td>
<td>631</td>
<td>68</td>
<td>636</td>
<td>6.3</td>
<td>26.1</td>
</tr>
<tr>
<td>Flamm</td>
<td>202</td>
<td>201</td>
<td>Experienced without NR</td>
<td>20</td>
<td>113</td>
<td>33</td>
<td>140</td>
<td>na</td>
<td>28.3</td>
</tr>
<tr>
<td>Prove-1</td>
<td>263</td>
<td>250</td>
<td>Naïve</td>
<td>27</td>
<td>160</td>
<td>0</td>
<td>157</td>
<td>6.6</td>
<td>26.7</td>
</tr>
<tr>
<td>Prove-2</td>
<td>256*</td>
<td>245</td>
<td>Naïve</td>
<td>5</td>
<td>103</td>
<td>1</td>
<td>149</td>
<td>6.4</td>
<td>23.7</td>
</tr>
<tr>
<td>Prove-3</td>
<td>354**</td>
<td>342</td>
<td>Experienced with NR</td>
<td>30</td>
<td>201</td>
<td>52</td>
<td>234</td>
<td>6.7</td>
<td>27.7</td>
</tr>
<tr>
<td>Realize</td>
<td>663</td>
<td>662</td>
<td>Experienced with NR</td>
<td>30</td>
<td>298</td>
<td>169</td>
<td>460</td>
<td>6.6</td>
<td>27.4</td>
</tr>
<tr>
<td>Respond-2</td>
<td>404</td>
<td>403</td>
<td>Experienced no NR</td>
<td>49</td>
<td>236</td>
<td>49</td>
<td>268</td>
<td>Na</td>
<td>28.3</td>
</tr>
<tr>
<td>Sprint-1</td>
<td>598</td>
<td>520</td>
<td>Naïve</td>
<td>78</td>
<td>288</td>
<td>37</td>
<td>305</td>
<td>6.6</td>
<td>Na</td>
</tr>
<tr>
<td>Sprint-2</td>
<td>1099</td>
<td>1097</td>
<td>Naïve</td>
<td>159</td>
<td>698</td>
<td>48</td>
<td>656</td>
<td>6.5</td>
<td>27.7</td>
</tr>
<tr>
<td>Tibotec-Japan</td>
<td>189</td>
<td>189</td>
<td>Naïve</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>99</td>
<td>6.8</td>
<td>22.8</td>
</tr>
</tbody>
</table>

N = number; % = proportion; Gen 1ª = HCV genotype 1ª infected patients; Cirr. = patients with cirrhosis; *this does not include the 78 patients included in the ribavirin-free arm; **this does not include the 111 patients included in the ribavirin-free arm.
<table>
<thead>
<tr>
<th>Study features</th>
<th>SVR</th>
<th>SAE</th>
<th>Severe anaemia</th>
<th>Severe infections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study ID</td>
<td>RR (CI95%)</td>
<td>n</td>
<td>RR (CI95%)</td>
<td>n</td>
</tr>
<tr>
<td>Advance PR48</td>
<td>0.5 (0.4-0.6)</td>
<td>361</td>
<td>0.4 (0.3-0.5)</td>
<td>24</td>
</tr>
<tr>
<td>T8PR-(RGT 24-48)</td>
<td>0.5 (0.4-0.6)</td>
<td>250</td>
<td>0.5 (0.4-0.6)</td>
<td>31</td>
</tr>
<tr>
<td>T12PR-(RGT 24-48)</td>
<td>0.5 (0.4-0.6)</td>
<td>182</td>
<td>0.5 (0.4-0.6)</td>
<td>24</td>
</tr>
<tr>
<td>Flamm PR48</td>
<td>0.5 (0.4-0.6)</td>
<td>134</td>
<td>0.5 (0.4-0.6)</td>
<td>14</td>
</tr>
<tr>
<td>Japan T12PR24</td>
<td>0.5 (0.4-0.6)</td>
<td>126</td>
<td>0.5 (0.4-0.6)</td>
<td>61</td>
</tr>
<tr>
<td>Prove-1 T12PR12</td>
<td>0.5 (0.4-0.6)</td>
<td>75</td>
<td>0.5 (0.4-0.6)</td>
<td>31</td>
</tr>
<tr>
<td>Prove-2 T12PR12</td>
<td>0.5 (0.4-0.6)</td>
<td>82</td>
<td>0.5 (0.4-0.6)</td>
<td>49</td>
</tr>
<tr>
<td>Prove-3 T12PR24</td>
<td>0.5 (0.4-0.6)</td>
<td>114</td>
<td>0.5 (0.4-0.6)</td>
<td>57</td>
</tr>
<tr>
<td>Realize T12PR48</td>
<td>0.5 (0.4-0.6)</td>
<td>264</td>
<td>0.5 (0.4-0.6)</td>
<td>175</td>
</tr>
<tr>
<td>Sprint-1 T12PR28</td>
<td>0.5 (0.4-0.6)</td>
<td>104</td>
<td>0.5 (0.4-0.6)</td>
<td>58</td>
</tr>
<tr>
<td>Sprint-2 T12PR48</td>
<td>0.5 (0.4-0.6)</td>
<td>363</td>
<td>0.5 (0.4-0.6)</td>
<td>171</td>
</tr>
</tbody>
</table>

**Note:** RR: risk ratio; CI95%: confidence interval; n: number of events.
Boceprevir and telaprevir for hepatitis C

The inverse variance model indicates that triple therapy with either boceprevir or telaprevir improved the SVR by 76% (RR 1.76 CI-95% 1.63-1.89). However, heterogeneity was strong (I-2 79.7% p for heterogeneity <0.001). SAEs (Figure 2B) occurred in 382 out of the 3,140 patients who received DAA, and in 105 out of the 1,337 patients who received SoC. The inverse variance model indicates that triple therapy with either boceprevir or telaprevir increased the risk of SAE by 52% (RR 1.52 CI-95% 1.23-1.88) with no detectable heterogeneity. Severe anaemia (Figure 2C) occurred in 160 out of the 3,140 patients who received DAA, and in 23 out of the 1,337 patients who received SoC. The inverse variance model indicates that triple therapy with either boceprevir or telaprevir increased the risk of severe anaemia more than doubled the risk of severe anaemia (RR 2.29 CI-95% 1.49-3.52) with no detectable heterogeneity. Severe infections (Figure 2D) occurred in 104 out of the 3,140 patients who received DAA, and in 22 out of the 1,337 patients who received SoC. The inverse variance model indicates that triple therapy with either boceprevir or telaprevir increased the risk of severe infections by 87% (RR 1.87 CI-95% 1.19-2.95) with no detectable heterogeneity. The most frequent clinical presentations of infections are reported Table 3.

Across-study risk of bias

No evidence of a potential small study effect was found. Harbord’s test p-value for no small study effect was 0.278, 0.661, 0.280 and 0.584 in the meta-analyses of SVR, SAE, severe anaemia and severe infection, respectively.
TABLE 3 - Most frequent sites of infection among 3,140 patients receiving triple therapy with direct-acting antiviral (DAA) and 1,337 receiving SoC.

<table>
<thead>
<tr>
<th>Clinical presentation</th>
<th>Receiving DAA (n=3140)</th>
<th>Not receiving DAA (n=1337)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRTI</td>
<td>28 (26.92)</td>
<td>5 (22.73)</td>
</tr>
<tr>
<td>BSI</td>
<td>11 (10.58)</td>
<td>1 (4.55)</td>
</tr>
<tr>
<td>SSTI</td>
<td>26 (25.00)</td>
<td>3 (13.64)</td>
</tr>
<tr>
<td>GI</td>
<td>17 (16.35)</td>
<td>5 (22.73)</td>
</tr>
<tr>
<td>URTI</td>
<td>5 (4.81)</td>
<td>2 (9.09)</td>
</tr>
<tr>
<td>UTI</td>
<td>6 (5.77)</td>
<td>3 (13.64)</td>
</tr>
<tr>
<td>IAI</td>
<td>9 (8.65)</td>
<td>3 (13.64)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (1.92)</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>104 (100)</td>
<td>22 (100)</td>
</tr>
</tbody>
</table>

LRTI: lower respiratory tract infections; BSI: blood stream infections; SSTI: skin and soft tissue infections; GI: gastrointestinal infections; URTI: upper respiratory tract infections; UTI: urinary tract infections; IAI: intra-abdominal infections; Other: 1 post-procedural infection, 1 tooth abscess.

**Additional analyses**

**Analyses of heterogeneity**

Sub-group analyses to explain heterogeneity in SVR meta-analysis indicated that response to previous treatment and cirrhosis at baseline may be potential causes of heterogeneity. In particular, response to previous treatment can fully explain heterogeneity (Figure 3A) while sub-group analysis for populations including more than 1% of cirrhotic participants can only partially explain the heterogeneity (Figure 3B). Simple meta-regression models provide evidence for a positive log-linear association between RR for SVR and either proportion of male patients (Figure 3C) or population mean BMI (Figure 3D). No evidence of association was present for: the proportion of HCV genotype 1a (studies included 10; p for linear association 0.537; I²=81.62%), the proportion of black patients (studies included 10; p for linear association 0.882; I²=81.80%), population mean age (studies included 10; p for linear association 0.124; I²=80.88%) and population mean log HCV-RNA at baseline (studies included 8; p for linear association 0.474; I²=80.88%). The multivariate meta-regression model indicated that the response to previous treatment is the only covariate independently associated to SVR, which can also entirely explain all the heterogeneity (Table 4).

**Sensitivity analyses**

Sensitivity analysis confirms robustness of evidence for all meta-analyses but the risk of severe anaemia. In particular, meta-analyses only including studies without risk of potential bias due to attrition indicate RR increases by about 22% in the restricted analyses in comparison to the overall analysis (9 studies RR 2.29 and p<0.001; 4 studies pooled RR 2.81 and p<0.020).

**Meta-analyses of dose response**

The analysis comprises a subset of the data consisting in the studies which evaluated different exposure to DAA plus a fixed 48 weeks of peg-IFN plus ribavirin. Sufficient data were only available for telaprevir, as boceprevir schemes in combination with a fixed 48 weeks of peg-IFN plus ribavirin are always very similar: either 44 or 48 weeks. The analysis included 7 arms from 3 different RCTs (i.e. PROVE1, arms PR48 and T12PR48; PROVE3, arm PR48 and T24PR48; REALIZE, arms PR48 T12PR48 with led-in, T12PR48 without led-in) for a total of 1,043 patients with 3 actual degrees of exposure to DAA (i.e. 0, 12 and 24 weeks). Overall, the analysis shows that both SVR and risk for SAE recognized a significant positive log-linear association with the exposure to DAA. In particular, SVR increased 3% per week of treatment in naïve patients who received telaprevir for 0-24 weeks (RR 1.03 95% CI 1.01-1.07; P-value 0.007). The increase was estimated to be about 2.28 times higher in treatment-experienced patients (RR 2.28 95%CI 1.40-3.72; P-value <0.001). Dose-response analyses for SAE were carried out without correction for previous response to therapy, as no heterogeneity was found. In this case, the model estimated that the risk of SAE increased about 5% per week of treatment (RR 1.05 95%CI 1.02-1.08; P-value <0.001).

**Overall quality of the body of evidence**

Figure 4 shows the diagram reporting the results of GRADE assessment and Table 5 reports the overall summary of evidence.
DISCUSSION

This study is a systematic review and meta-analysis of evidence from all currently available RCTs comparing triple therapy with DAA to SoC. Several issues have emerged. Firstly, triple therapy produces a higher proportion of SVR than SoC. The inverse variance model estimates that SVR is increased by 1.76 times in patients receiving DAA, though the heterogeneity is significant. Additional analyses to explore the potential causes of heterogeneity suggest that the patients’ response to previous treatment acts as the strongest effect modifier.

![Data Table and Graphs]

**FIGURE 3** - Analyses of heterogeneity in meta-analyses for SVR through sub-group analyses and simple meta-regression model. A: Sub-group analyses for previous response to interferon therapy. B: Sub-group analyses for having included more than 1% of cirrhotic patients. C: simple meta-regression model according to proportion of males (p for linear association 0.003; I2=51.37%). D: simple meta-regression model according to population mean baseline BMI. (RR: risk ratio; 95% CI: 95% confidence interval; I-V: inverse variance fixed effect model; D+L: DerSimonian and Laird random effect model; WT%: proportion weight in the I-V model; Ln: natural logarithm; BMI: mean body max index.

**TABLE 4** - Multiple meta-regression model to assess the potential cause of heterogeneity. Only 9 studies were included, as SPRINT-I does not report data about mean BMI. The analyses show that previous response to interferon therapy alone can explain all the heterogeneity found in the inverse variance model for SVR.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Num. of studies</th>
<th>Coef. (95% CI)</th>
<th>p-value</th>
<th>I2 Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response to previous treatment</td>
<td>9</td>
<td>1.57 (1.08-2.29)</td>
<td>0.030</td>
<td></td>
</tr>
<tr>
<td>Inclusion of cirrhotic patients</td>
<td></td>
<td>1.08 (0.68-1.70)</td>
<td>0.670</td>
<td></td>
</tr>
<tr>
<td>Population mean BMI</td>
<td></td>
<td>1.03 (0.88-1.20)</td>
<td>0.623</td>
<td>0.00%</td>
</tr>
<tr>
<td>Male proportion in the population</td>
<td></td>
<td>1.00 (0.93-1.06)</td>
<td>0.863</td>
<td></td>
</tr>
</tbody>
</table>
which can explain all the heterogeneity. In fact, the effect of DAA is much greater in the RCTs including experienced patients than in those enrolling treatment-naive patients only. The finding that the relative effect of DAA is greater in difficult-to-treat patients is well-known, and re-analyses of RCTs have already pointed out that the advantage of triple therapy over SoC is attenuated, though present, in patients with the favourable IL-28B CC genotype (Poordad et al., 2012; Jacobson et al., 2011). Secondly, ribavirin is not associated with an improvement in SVR or with a significant reduction of adverse events. However, this observation is tentative as:

a) only a small number of patients received ribavirin-free therapy (i.e. 2 RCTs with telaprevir only);

b) this observation relies on the descriptive analysis, as we did not include ribavirin-free regimens in the inferential analysis.

Thirdly, the study provides evidence that triple therapy increases the incidence of severe adverse events. In particular, the risk of experiencing either any SAE, severe anaemia or severe infection increases by 1.52, 2.29 and 1.87, respectively. Severe anaemia is the only outcome for which the DAA effect significantly changes in the sensitivity analysis. This may be due to the fact that restricted analysis refers only to telaprevir studies which, in contrast to boceprevir RCTs, did not allow for erythropoietin. Severe infections occurred with a cumulative 3% incidence across the study. Although they are infrequent events, the evidence that DAA can increase the risk of severe infections suggests that:

a) the increased proportion of leukopenia due to DAA (Pearlman, 2012) has an actual clinical consequence;

b) the currently ongoing RCTs to assess the efficacy of DAA in special population, such as HIV-positive subjects and solid transplant recipients, should pay particular attention to the issue of severe infections.

Fourthly, we found that both SVR and SAEs have a dose response relation with telaprevir. This strengthens the overall quality of evidence for the causality link between triple therapy and either SVR or SAEs. Unfortunately, the same analyses could not be made for boceprevir due to lack of specific data and the different regimen schemes for this compound.

Finally, currently available data provide no solid evidence that one DAA is superior to the other. In fact, no residual heterogeneity is present in subgroup analyses for SVR and none was found for adverse events. This evidence conflicts with recently published results from several other studies. In particular, at least 6 meta-analyses comparing boceprevir with telaprevir have been published over the last 12 months.
Three of these used a frequentist approach with a random effect model and indirect comparison (Chou R et al., 2013; Cooper et al., 2012; Sitole et al., 2013) and 3 are Bayesian meta-analyses (Cooper et al., 2013; Kieran et al., 2013; Cure et al., 2012). None used a frequentist approach with extended analyses of the potential causes of heterogeneity, as in our study. The frequentist meta-analyses indicate that boceprevir and telaprevir perform substantially equally, although one found that telaprevir is marginally better than boceprevir with regard to SVR (i.e. superior at 24 weeks but not at the 48 week follow-up) (Sitole et al., 2013). In contrast, the three Bayesian meta-analyses provided more divergent results. The first provides evidence that the two DAA always perform equally (Cooper et al., 2013), the second indicates that telaprevir is better than boceprevir in a subset of treatment-experienced patients only (Kieran et al., 2013), while the third suggests that telaprevir is always superior to boceprevir (Cure et al., 2012). It might be worthy of notice that the DAA manufacturer directly funded two of the three Bayesian meta-analyses but none of the frequentist meta-analyses.

Bayesian approaches are endowed with remarkable flexibility in fitting models for meta-analyses of multiple treatment RCTs, but their predictions strongly depend on the consistency of the a priori distribution assumption and reliability of the assumption that participants across the studies are similar in all but the exposure to the intervention to be assessed. In fact, Bayesian meta-analyses have been found to be exposed to major misinterpretation when used to pool results from heterogeneous RCTs with complex study designs, (Thijs et al., 2008; Diener et al., 2008; Lam et al., 2007; Lam et al., 2008) such as in the present circumstance. We believe that different a priori distribution assumptions may explain inconsistencies between Bayesian meta-analyses, while the assumption of the population homogeneity across the studies may explain the difference from our results. Indeed, recently published results from the PROVIDE study suggest that either DAA has similar effects even in null-responders; i.e. 38% of SVR in the PROVIDE study (Vierling et al., 2011) and 41% in the REALIZE study. (Zeuzem et al., 2011).

Major caution should be exercised when interpreting the results. Firstly, the observation that previous response to treatment is the only co-

### TABLE 5 - Summary of evidence.

<table>
<thead>
<tr>
<th>Meta-analysis</th>
<th>RR (95% CI)</th>
<th>p for RR=1</th>
<th>Number of studies</th>
<th>Number of participants</th>
<th>Quality of evidence</th>
<th>Heterogeneity:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sustained virological response</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>1.76 (1.63-1.89)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td>None</td>
<td>H strong in overall analyses (I²= 79.7%) but completely explained by sub-group analysis according to the response to previous treatment</td>
</tr>
<tr>
<td>Naïve</td>
<td>1.62 (1.50-1.74)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Experienced including NR</td>
<td>3.84 (2.85-5.17)</td>
<td>&lt;0.001</td>
<td>10</td>
<td>4997</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Experienced without NR</td>
<td>3.00 (2.18-4.14)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Severe adverse reactions</td>
<td>1.52 (1.23-1.88)</td>
<td>&lt;0.001</td>
<td>9</td>
<td>4477</td>
<td>Low</td>
<td>None</td>
</tr>
<tr>
<td>Severe anaemia</td>
<td>2.29 (1.49-3.52)</td>
<td>&lt;0.001</td>
<td>9</td>
<td>4477</td>
<td>Moderate</td>
<td>None</td>
</tr>
<tr>
<td>Severe infections</td>
<td>1.74 (1.05-2.87)</td>
<td>0.031</td>
<td>9</td>
<td>4477</td>
<td>Low</td>
<td>None</td>
</tr>
</tbody>
</table>

NR: non-responder; H: heterogeneity; RR: risk ratio; 95%CI: 95% confidence interval.
variate exerting a significant interaction does not mean that, at the individual patient level, other baseline characteristics (e.g. IL-28, sex, stage of liver diseases and BMI) have no effect on SVR. In fact, meta-analyses based on aggregate data fail to detect potentially relevant effect modifiers even when they are even grossly homogeneous across the studies and/or are associated with a stronger interaction term. Secondly, pooled RR for SVR is estimated through a sub-group analysis which in principle may suffer from the same limits as observational studies. Nevertheless, all the heterogeneity we found is quantitative (i.e. it affects the size of the effect but not the directions), therefore the random-effect model for SVR still provides a good average of DAA effect based on randomized observations.

Thirdly, potential imprecision in the estimate of pooled RRs for SAEs, severe anaemia and severe infections can be present as they are supported by meta-analyses which do not meet OIS criteria; this is also why the overall body of evidence for these outcomes was downgraded to moderate and low quality. Fourthly, we had to aggregate the observations of different intervention arms within the same RCTs to obtain individual pair-wise comparisons to be analysed, and this may have potentially distorted the punctual estimate for RRs. Finally, as with many meta-analyses of RCTs, our study includes RCTs conducted in clinical centres in rich countries (mainly Europe, North America and Japan), so our results cannot be directly generalized to settings with significantly different features (e.g. low- and middle-income settings in Africa and Asia).

Despite the above-mentioned limitations, this study provides solid evidence that triple therapy can remarkably improve the proportion of SVR in treatment-experienced and treatment-naive patients. However, meta-analyses on adverse events provide straightforward evidence that triple therapy increases the incidence of overall SAEs, severe anaemia and severe infections. In particular, evidence of severe infections is, in principle, free from the effect of potential confounders, as it comes from fixed-effect model meta-analyses with no heterogeneity thereby preserving the effect of randomization in the primary studies. It is noteworthy that the efficacy of boceprevir and telaprevir is currently investigated in special groups such as HCV/HIV co-infected subjects and solid organ transplant recipients who are in principle at higher risk of both severe infection and anaemia due to baseline immunodepression and drug interactions. Our study provides preliminary information for planning new clinical studies and warning doctors and clinical researchers of the potential risk of severe unexpected events of therapy with first generation protease inhibitors in particular groups of patients.

**List of abbreviations in alphabetical order**

- BMI: body mass index
- CI-95%: 95% confidence interval
- DAA: direct-acting antiviral
- HCV: Hepatitis C virus
- I-2: I squared statistic
- NCT: National Clinical Trial Registry
- Peg-IFN: pegylated interferons
- RCT: randomized controlled trial
- RGT: response guided therapy
- RNA: ribonucleic acid
- RR: risk ratio
- SAE: severe adverse event
- SoC: standard of care
- SVR: sustained virological response

**Competing interest declaration**

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare that:

1) none of them have support for the submitted work;
2) and/or have or have had relationships with companies that might have an interest in the submitted work in the previous 3 years;
3) and/or their spouses, partners or children have no financial relationships that may be relevant to the submitted work;
4) and/or have any non-financial interests that may be relevant to the submitted work.

**Authors’ contributions**

SL designed the study, set up the databases for data collection, defined the search strings, performed the literature search, selected and assessed papers, extracted data, made the analysis, interpreted results and drafted the text. AM performed the literature search, selected and
assessed papers, extracted data, implemented non-standard meta-analyses techniques, interpreted the results, and reviewed and approved the final text. VP, EG, BP and GI interpreted the results, reviewed and approved the final text. GI approved the study design, interpreted the results, reviewed the text and gave final approval to the paper.

ACKNOWLEDGEMENTS
The study was funded by the Italian Ministry of Health “Ricerca Corrente 2012”. The Ministry had no role in the study. All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and accuracy of the data analysis.

List of the studies included in the meta-analysis: in bold underlined the study ID as reported in the paper in brackets "[ ]” the number of registered clinical trial available at https://clinicaltrials.gov


REFERENCES


