Comparative Efficacy and Tolerability of Janus Kinase Inhibitor Therapies for Moderate to Severe Crohn’s Disease: A Network Meta-analysis
Theodore Rokkas¹², Konstantinos Ekmektzoglou¹², Yaron Niv³

INTRODUCTION

Crohn's disease (CD) is a chronic inflammatory bowel disease (IBD), characterized by relapses and progressive intestinal lesions which may result in complications such as abscesses, fistulae and strictures [1, 2]. Over recent decades, biologic therapies, i.e. against tumor necrosis factor-alpha (TNF-a), leucocyte integrins and interleukin (IL) pathways, have become significant therapies in our armamentarium for moderate-to-severe CD and ulcerative colitis (UC) [3]. However, there are some limitations accompanying these therapies as a percentage of patients do not respond and, in some, there is a loss of response over time [4, 5]. In addition, currently available biologic treatments are administered intravenously or subcutaneously, which potentially represents a burden for patients and might compromise their compliance to treatment [6]. Furthermore, despite the introduction of lower-cost biosimilars, the cost of biologic treatments remains an important challenge for healthcare budgets [7]. Finally, there are some safety issues which are related with both traditional and biologic treatments [8-10]. For all the above reasons, novel treatment options easily administered, effective, well-tolerated and safe are required.

ABSTRACT

Background & Aims: A number of Janus kinase (JAK) inhibitors (tofacitinib, filgotinib, upadacitinib) have been tested for moderate and severe Crohn's disease (CD) in randomized control trials (RCTs). However, data on their comparative efficacy and tolerability is lacking. We aimed to study their performance comparatively, by means of network meta-analysis (NWM).

Methods: We searched the Pubmed/Medline, EMBASE, and Cochrane Library databases for relevant RCTs through March 2021 and data was extracted. A bayesian NWM was performed to investigate the efficacy and tolerability of the above JAK inhibitors and to explore their rank order in treating moderate and severe CD patients. The cumulative ranking probability for each intervention at the end of treatment period, was evaluated by means of surfaces under cumulative ranking (SUCRA) values.

Results: Four RCTs were entered into this NWM. They included 811 patients totally, randomized to 11 interventions, i.e. placebo, tofacitinib (1mg BID, 5mg BID, 10mg BID, 15mg BID), filgotinib 200 OD and upadacitinib (3 mg BID, 6 mg BID, 12 mg BID, 24 mg BID and 24mg OD). Two upadacitinib doses (6 mg BID and 24 mg BID) and filgotinib 200 OD, performed best as judged by the relevant forest plots, league matrices, rankograms, SUCRA values (96.7%, 84.6 %and 78.7%, respectively) and the clustered ranking plots for efficacy and tolerability.

Conclusions: Upadacitinib 6 mg BID, upadacitinib 24 mg BID and filgotinib 200 OD performed better as induction therapies in comparison to control therapies. Consequently, these regimens may play a therapeutic role in CD and therefore they merit further evaluation with well-designed RCTs.

Key words: Crohn's disease - JAK inhibitors - efficacy – tolerability - network meta-analysis.

Abbreviations: BID: twice daily; CD: Crohn's disease; CDAI: Crohn's disease activity index; CI: confidence interval; CrI: credible interval; IBD: inflammatory bowel disease; IL: interleukin; JAK: Janus kinase; NWM: network meta-analysis; OD: once daily; OR: odds ratio; RCT: randomized controlled trial; RoB: risk of bias; SUCRA: surfaces under cumulative ranking; TNF: tumor necrosis factor; UC: ulcerative colitis.

Received: 02.06.2021
Accepted: 08.08.2021

Address for correspondence: Theodore Rokkas MD, PhD, Gastroenterology Clinic, Henry Dunant Hospital, 107 Messogion Ave., Athens 11526, Greece sakkor@otenet.gr
Janus kinase (JAK) inhibitors are orally administered, non-immunogenic small molecule drugs. Four intracellular tyrosine kinase (TYK) proteins, i.e. JAK1, JAK2, JAK3 and TYK2, are included in the JAK group and their activation is responsible for initiating the intracellular signaling associated with different cytokine receptors. JAK phosphorylation subsequently activates the intracytoplasmic signal transducer and activator of transcription (STAT) pathways to control downstream target gene expression of inflammatory mediators [11,12]. Consequently, research has been focused on the clinical application of JAK inhibitors in immune-mediated inflammatory diseases and among other indications, considerable attention is currently being paid to the clinical use of JAK inhibitors in IBD. Three JAK inhibitors (tocafitinib, upadacitinib and filgotinib) are now under consideration for use in IBD and of those tocafitinib has been approved for oral use in moderate-to-severe UC [13]. In parallel, there are RCTs comparing these JAK inhibitors to placebo for moderate-to-severe CD. However, knowledge of their comparative efficacy and tolerability is lacking.

The comparative efficacy and tolerability of RCTs, concerning multiple treatments competing for a similar therapeutic result, can be achieved by utilizing a useful evidence synthesis tool, i.e. a network meta-analysis (NWM) [14-16]. Network meta-analysis incorporates both direct and indirect evidence, thus providing information concerning the relative effects of treatments included in relevant RCTs. The comparative efficacy and tolerability of JAK inhibitors for moderate-to-severe CD has not been explored. In this study, therefore, we aimed to examine the above, by means of NWM of published RCTs.

METHODS

To identify studies and extract data in this NWM, we have followed the steps (i.e. identification, screening, eligibility, inclusion) described in our previous publications [17]. Thus, the PubMed/MEDLINE and Embase databases were searched until March 2021 to identify human studies written in English using the following search text and/or Medical Topic Heading (MeSH) terms: (”janus kinase inhibitors”[All Fields] OR ”janus kinase inhibitors”[MeSH Terms] OR (”janus”[All Fields] AND ”kinase”[All Fields] AND ”inhibitors”[All Fields]) OR ”janus kinase inhibitors”[All Fields]) AND (IBD[All Fields] OR (”crohn disease”[MeSH Terms] OR (”crohn”[All Fields] AND ”disease”[All Fields]) OR (”crohn disease”[All Fields] OR ”crohn disease”[All Fields] OR ”crohn’s disease”[All Fields]) OR ”crohn’s disease”[All Fields])). In addition, a manual search of all review articles, published editorials and retrieved original studies, was made. Two authors (T.R and K.E) independently extracted data from each study. Any disagreement was settled with further discussion until consensus was reached. This NWM was performed according to the PRISMA statement for interventions [18], whereas the rating of the quality of treatment effect estimates was achieved by using the GRADE (i.e. Grading of Recommendations Assessment, Development and Evaluation) working group modality [19]. Furthermore, we appraised the confidence in estimates derived from this NWM, as described in our previous publications [20, 21].

We defined the inclusion and exclusion criteria before starting the study investigation. Thus, appropriate studies were included provided that the following criteria were met a) published as complete articles or abstracts with data that could be extracted; b) written in English, and c) RCTs with JAK inhibitors in one arm. Studies not meeting the above criteria were excluded. In this NWM, the induction of clinical remission, i.e. Crohn’s disease activity index (CDAI) < 150, was defined as the end point for efficacy. For RCTs reporting on maintenance of induced remission, outcomes were assessed at the last point of follow-up.

For pair-based meta-analyses and heterogeneity estimation (Cochran’s Q test and the I² metric), we followed the methodology described previously [17]. In addition to heterogeneity, we assessed inconsistency, i.e. the agreement between direct and indirect evidence and the transitivity assumption, as these are critical when conducting an NWM [22, 23]. We constructed comparison-adjusted funnel plots and checked their symmetry to assess whether small-scale trials influence the efficacy results. Surfaces under cumulative ranking (SUCRA) values were used in intervention network charts to examine the cumulative ranking probability for each intervention concerning the efficacy achieved by this intervention compared to an ideal intervention showing the best efficacy without doubt, i.e. SUCRA = 1 or 100% when expressed as a percentage [14-16]. Except for efficacy, tolerability was taken into account and all the competing treatments therefore were compared and ranked hierarchically according to their performance on two outcomes, i.e. efficacy and tolerability. We achieved this by constructing a two-dimensional clustered ranking plot with the relevant dendrogram, presenting jointly the relative ranking of treatments (based on SUCRA values) for efficacy and tolerability. Using this hierarchical method, we were able to detect clusters of treatments with similar performance on both outcomes [14-16]. Data were processed using software suitable for bayesian network meta-analysis, namely Stata 13.2 (StataCorp, College Station, TX) [14, 15] and NetMetaX [16]. In all included RCTs, the intention to treat results (ITT) were taken into account.

RESULTS

Characteristics of Studies

The process of study selection is shown in Fig. 1. Thus, out of 977 titles yielded by the initial search, 4 RCTs were eligible for meta-analysis [24-27]. The characteristics of these RCTs are shown in Table I. They were phase II, double-blind, randomized, placebo-controlled trials, in patients with moderate-to-severe CD and included 811 patients totally, randomized to 11 treatments, i.e. placebo, tofacitinib 1 mg BID, tofacitinib 5 mg BID, tofacitinib 10 mg BID, tofacitinib 15 mg BID, upadacitinib 3 mg BID, upadacitinib 6 mg BID, upadacitinib 12 mg BID, upadacitinib 24 mg BID and filgotinib 200 mg BID. The RCT by Sandborn et al. [24], included 4 arms, i.e. placebo, tofacitinib 1 mg BID, 5 mg BID and 15 mg BID. The RCT by Panés et al. [25], included 4 arms, i.e. placebo, tofacitinib 5 mg BID, tofacitinib 10 mg BID and tofacitinib 15 mg BID. The RCT by Vermeire et al. [26], included 2 arms, i.e. placebo and filgotinib 200 mg OD. Finally, the RCT by Sandborn et al. [27], included 6 arms, i.e. placebo, upadacitinib 3 mg BID, upadacitinib 6 mg BID, upadacitinib
12mg BID, upadacitinib 24mg OD and upadacitinib 24mg BID. Concerning quality assessment, the bar graph of Fig. 2A depicts the summary of the risk of bias (RoB), characterized as high, unclear and low RoB, concerning 5 items, i.e. random sequence generation (selection bias), allocation concealment (selection bias), incomplete outcome data (attrition bias), blinding of participants and personnel (performance bias) and publication format.

**Network Meta-analysis**

**Efficacy network map**

The network map of all 11 therapeutic interventions is shown in Fig. 3, with all 55 possible comparisons, i.e. 25 direct and 30 indirect. In this map the node size reflects the number of patients allocated to each treatment, whereas the edge thickness is in proportion to the precision, i.e. the inverse of variance of each direct comparison [16].

**Efficacy network forest and funnel plots**

The pair-wise comparisons [ORs (95% CI)] of all 25 direct treatment comparisons is depicted in the forest plot of Fig. 4. Out of these, 8 yielded significant results, i.e. upadacitinib 6 mg BID vs placebo (OR=6.38, 95%CI: 2.32-17.55), upadacitinib 6 mg BID vs upadacitinib 12 mg BID (OR=6.14, 95%CI: 2.22-16.92), upadacitinib 6 mg BID vs upadacitinib 3 mg BID (OR=5.40, 95%CI: 2.02-14.46), upadacitinib 6 mg BID vs upadacitinib 24 mg OD (OR=4.05; 95%CI: 1.50-10.92), upadacitinib 24 mg BID vs placebo (OR=3.71, 95%CI: 1.40-9.82), upadacitinib 24 mg BID vs upadacitinib 3 mg BID (OR=3.57; 95%CI: 1.35-9.47), upadacitinib 24 mg BID vs upadacitinib 3 mg BID (OR=3.14; 95%CI: 1.22-8.08) and filgotinib 200 mg OD vs placebo (OR=3, 95%CI: 1.37-6.58). There was no significant heterogeneity (Q=14.37, I²=23.47%, p=0.21). The respective comparison-adjusted funnel plot is shown in Fig. 2B. It appears symmetrical, implying the
absence of publication bias. All 55 possible comparisons (25 direct and 30 indirect) in this NWM are shown in the forest plot of Fig. 5A. Of these comparisons [ORs, 95% Credible Intervals (CrI)], 13 yielded significant results. In this network forest plot there was no significant heterogeneity and also the evaluation of inconsistency yielded insignificant overall results, meaning that the comparative effect sizes that were obtained by direct and indirect comparisons were consistent (Supplementary Table 1). When taking placebo as reference treatment, the results showed that, among 10 treatments tested against it, only three, i.e. upadacitinib 6 mg BID, upadacitinib 24 mg BID and filgotinib 200 mg OD yielded significant results (Fig. 5B). The remaining seven comparisons yielded insignificant results. None of the tofacitinid tested doses were superior to placebo.

### Efficacy league matrixes, rankograms and SUCRA values

The comparative efficacies [ORs (95% CrI)] of the 11 treatments is shown in the league matrix of Fig. 6A. The respective rankogram is shown in Fig. 6B in close relationship with SUCRA values shown in Table II. These results showed that upadacitinib 6 mg BID (SUCRA 96.7%), upadacitinib 24 mg BID (SUCRA 84.6%) and filgotinib 200 OD (SUCRA 78.7%) performed better in comparison to rest of the treatments.
Tolerability and Hierarchical Cluster Analysis

The included RCTs in the NWM did not evaluate tolerability. However, they reported withdrawals due to side effects and from this data tolerability could be inferred for the treatments tested. Hence an additional tolerability NWM was performed (Fig. 7). Taking into account SUCRAs for both efficacy and tolerability (Table II) we constructed the relevant clustered ranking plot for the treatments included in this NWM (Fig. 8A) with the respective dendrogram depicting the treatment similarities (Fig. 8B). Both these figures show that the 11 treatments formed 4 clusters with similar performance. One of those clusters, i.e. upadacitinib 6 mg, upadacitinib 24 mg BID and filgotinib 200 mg, showed the best combined performance concerning efficacy-tolerability and achieved the best similarity in the relevant dendrogram.

DISCUSSION

Janus kinase inhibitors have been approved for a variety of diseases, such as autoimmune diseases, rheumatoid arthritis...
and psoriasis [28, 29]. In IBD approval was given in 2018 to tofacitinib for the treatment of moderate-to-severe UC [30, 31], whereas for CD there are only published RCTs. In this NMA, we examined the comparative efficacy and tolerability of three JAK inhibitors, i.e. tofacitinib, filgotinib, upadacitinib, in inducing clinical remission in patients with moderate-to-severe CD, as assessed by the results of relevant RCTs. All four RCTs included in this NWM, were phase II trials. The results showed that, concerning efficacy, among 11 treatments, i.e. placebo, tofacitinib (1mg BID, 5mg BID, 10mg BID, 15mg BID), upadacitinib (3 mg BID, 6 mg BID, 12 mg BID, 24 mg BID, 24 mg OD) and filgotinib 200 OD, the cluster of three treatments, i.e. upadacitinib 6 mg BID, upadacitinib 24 mg BID and filgotinib 200 OD, performed best as judged by the pair-wise and network forest plots, SUCRA values and the relevant league matrixes and rankograms. Tofacitinib did not prove effective in this NWM and in fact, in its RCTs, placebo response was high, raising concerns for future research.

In addition to efficacy, we also performed a tolerability NWM calculating SUCRA values for each treatment and creating the relevant rankograms. Both efficacy and tolerability profiles were then taken into account in the created clustered ranking plot and dendrogram. Through this methodology we were able to assess treatment efficacy-tolerability profile and we found that the cluster of three treatments, i.e. upadacitinib 6 mg BID, upadacitinib 24 mg BID and filgotinib 200 OD, remained the best performer. All this might signal that these three treatments merit further evaluation concerning efficacy and tolerability in well-designed phase III RCTs in patients with moderate-to-severe CD. Towards this notion, plans are already going ahead for upadacitinib [27, 32].

In two of the included RCTs [25, 27], after the induction period, the authors re-randomized patients and provided data concerning maintenance assessment of efficacy and safety. Thus, Panés et al. [25], reported that clinical response was maintained at 26 weeks in 46.5% of CD patients treated with tofacitinib compared to 35.7% of patients treated with placebo. In addition, clinical remission was maintained in 39.5% of patients treated with tofacitinib compared to 28.5% of patients treated with placebo. In this study, the proportion of patients maintaining clinical response or remission, with either tofacitinib 5 mg BID or 10 mg BID, was not significantly different compared to placebo. Sandborn et al. [27], in their RCT evaluated upadacitinib as maintenance treatment in patients with moderate-to-severe CD and found that, after the induction period, efficacy was maintained for most endpoints through week 52. Concerning safety, the authors reported that over 52 weeks, the upadacitinib safety profile was consistent with studies in rheumatoid arthritis. However, patients in the twice-daily 12 mg and 24 mg upadacitinib groups had...
significant increases in total, high-density lipoprotein, and low-density lipoprotein cholesterol levels compared with patients in the placebo group. In general, the safety profile needs to be further investigated. Special attention should be given to thromboembolic events and herpes zoster, since an increased rate has been found in patients with UC and rheumatoid arthritis [33].

The approval of tofacitinib for the treatment of moderate-to-severe UC marked the first of several JAK inhibitors that are likely to enter the IBD treatment landscape within the next few years with significant implications for clinical practice. There are several reasons in favor of this hypothesis. Firstly, the oral bioavailability makes this treatment attractive, taking into account that all biologics in use are available as injections, with consequences concerning patients’ compliance. Furthermore, JAK inhibitor efficacy after anti-TNF failure and its lack of immunogenicity, make this treatment a potentially viable option for combination with biologics, such as vedolizumab, which has a favorable safety profile in patients with severe disease. Nevertheless, some unanswered questions remain. The fact that JAK inhibitors are oral therapies, does not necessarily mean that these drugs are safer than parenterally administered biologics. Furthermore, it is unclear whether JAK inhibitors will be more cost effective than currently available biosimilars. Finally, the efficiency of JAK treatment has to be improved by identifying individual patient predictors of response.

Fig. 6. A. SUCRA (surface under the cumulative ranking) based efficacy ranking league matrix showing the comparative efficacies of the treatments included in this network meta-analysis. B. Rankograms derived from relevant SUCRA values for the regimens evaluated in the included RCTs, showing the cumulative rank order for each intervention (1-11).
The good quality of the included RCTs, the lack of heterogeneity and publication bias and the lack of inconsistency strengthens the results of this NWM. However, some limitations should be mentioned. Thus, although heterogeneity was not observed for pooled efficacy or safety estimates, there was heterogeneity in the trial design, concerning treatment duration and endpoint definitions between trials that may have influenced the assessments. Furthermore, although potential risks associated with JAK inhibitors are likely to be dose-dependent, the small number of trials for each treatment precludes the accurate dose-response analysis. For all these reasons, more studies are needed. Thus, head-to-head RCTs are needed to evaluate the precise role of JAK inhibitors for the treatment of patients with moderate-to-severe CD together with biologics in use. Moreover, research is required to determine the exposure-response relationship of these drugs at the site of action, i.e. in the gut tissue. This might help the development of safer drugs when considering targeted therapies with less systemic exposure. An additional point of consideration would be to identify the suitable JAK inhibitor efficacy biomarkers and predictors which could guide CD patients more efficiently to the most effective therapy.

In summary, this NWM showed that JAK inhibitors could be an attractive therapeutic option in patients with moderate to severe CD, both naive and previously exposed to anti-TNF biologics with demonstrable efficacy and tolerability. This is strengthened when taking into account the fact that this class of drugs is administered orally. However, all included RCTs in this NWM were phase II studies. Therefore, carefully designed phase III RCTs are needed considering both efficacy and tolerability. In particular, carefully defined enrolment criteria with central reading of endoscopy for enrolment and endpoint adjudication can reduce biases. Additionally, further research can lead to predictors of response to JAK inhibitors. This is of importance when considering proper patients’ guidance to the most effective treatments.

Table II. SUCRA values for efficacy and tolerability. For efficacy high SUCRA denotes good efficacy. For tolerability high SUCRA denotes good tolerability

<table>
<thead>
<tr>
<th>TREATMENT</th>
<th>EFFICACY</th>
<th></th>
<th>TOLERABILITY</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUCRA (%)</td>
<td>PrBest</td>
<td>Mean Rank</td>
<td>SUCRA (%)</td>
</tr>
<tr>
<td>Placebo</td>
<td>20.1</td>
<td>0.0</td>
<td>9.0</td>
<td>41.8</td>
</tr>
<tr>
<td>Tofacitinib 1mg BID</td>
<td>60.5</td>
<td>2.4</td>
<td>5.0</td>
<td>37.6</td>
</tr>
<tr>
<td>Tofacitinib 5 mg BID</td>
<td>40.5</td>
<td>0.0</td>
<td>7.0</td>
<td>37.5</td>
</tr>
<tr>
<td>Tofacitinib 10 mg BID</td>
<td>37.5</td>
<td>0.0</td>
<td>7.2</td>
<td>65.0</td>
</tr>
<tr>
<td>Tofacitinib 15mg BID</td>
<td>21.8</td>
<td>0.0</td>
<td>8.8</td>
<td>29.7</td>
</tr>
<tr>
<td>Filgotinib 200 mg OD</td>
<td>78.7</td>
<td>9.9</td>
<td>3.1</td>
<td>47.7</td>
</tr>
<tr>
<td>Upadacitinib 3 mg BID</td>
<td>33.7</td>
<td>0.0</td>
<td>7.6</td>
<td>64.1</td>
</tr>
<tr>
<td>Upadacitinib 6 mg BID</td>
<td>96.7</td>
<td>76.0</td>
<td>1.3</td>
<td>47.9</td>
</tr>
<tr>
<td>Upadacitinib 12 mg BID</td>
<td>26.4</td>
<td>0.0</td>
<td>8.4</td>
<td>47.2</td>
</tr>
<tr>
<td>Upadacitinib 24 mg OD</td>
<td>49.5</td>
<td>0.0</td>
<td>6.0</td>
<td>57.0</td>
</tr>
<tr>
<td>Upadacitinib 24 mg BID</td>
<td>84.6</td>
<td>11.5</td>
<td>2.5</td>
<td>58.4</td>
</tr>
</tbody>
</table>
CONCLUSIONS

In this NWM, among 11 interventions tested as induction treatments in patients with moderate-to-severe CD, upadacitinib 6 mg BID, upadacitinib 24 mg BID and filgotinib 200 mg OD formed a cluster showing the best performance concerning efficacy and tolerability. However, additional well-designed RCTs are required to better understand the role of this class of drugs in the management of CD, alongside therapies in use.

Conflicts of interest: None to declare.

Authors’ contributions: T.R. conceived and designed this study, analyzed data and drafted the manuscript. K.E and Y.N. extracted and analyzed data, edited, and revised the manuscript. T.R. critically revised the paper. All the authors approved the final version of the manuscript.

Supplementary material: To access the supplementary material visit the online version of the J Gastrointestin Liver Dis at http://dx.doi.org/10.15403/jgld-3804

REFERENCES


