

SIMILAR

Can't see the forest for the trees? Plagiarism versus 'over-sensitive' similarity index

Sangita P Patil^a, Prasanthi Chengalva^a, Madhavi Patil^a and Vatsal Shah^b

^aSIRO Clinpharm Pvt. Ltd, Mumbai, India; ^bSIRO Clinpharm USA LLC, Princeton, NJ, USA

INTRODUCTION

- Similarity checking software in journal publishing can generate startling reports of similarity index, casting a doubt on the integrity of the publication.
- Passing the plagiarism check can be challenging, particularly, while developing a secondary manuscript, which has a previously published primary article or congress abstracts of the same study.
- It's imperative to understand the difference between text 'similarity' and plagiarism. The similarity index could capture innocent terms/phrases from the huge database of sources; such 'over-sensitive' detection can distract the focus from the major scientific content.

How a high similarity-index impacts us

- Shocking and **demotivating**.
- Irreparable consequences for agencies, academic authors and Publication Professionals and **damage reputation**.
- Dilemma of what and how much can be modified without compromising the integrity of the reported content.



OBJECTIVE

This case-study analyzed the similarity report for our article from a journal to objectively verify the extent of similarity and assess its validity for consideration as plagiarism.

METHODS



Phase 4 manuscript
(prior Phase 1-3 manuscripts /posters published)

Journal submission



Journal shared Similarity Index
(34%)

Analysis of individual sections of the manuscript



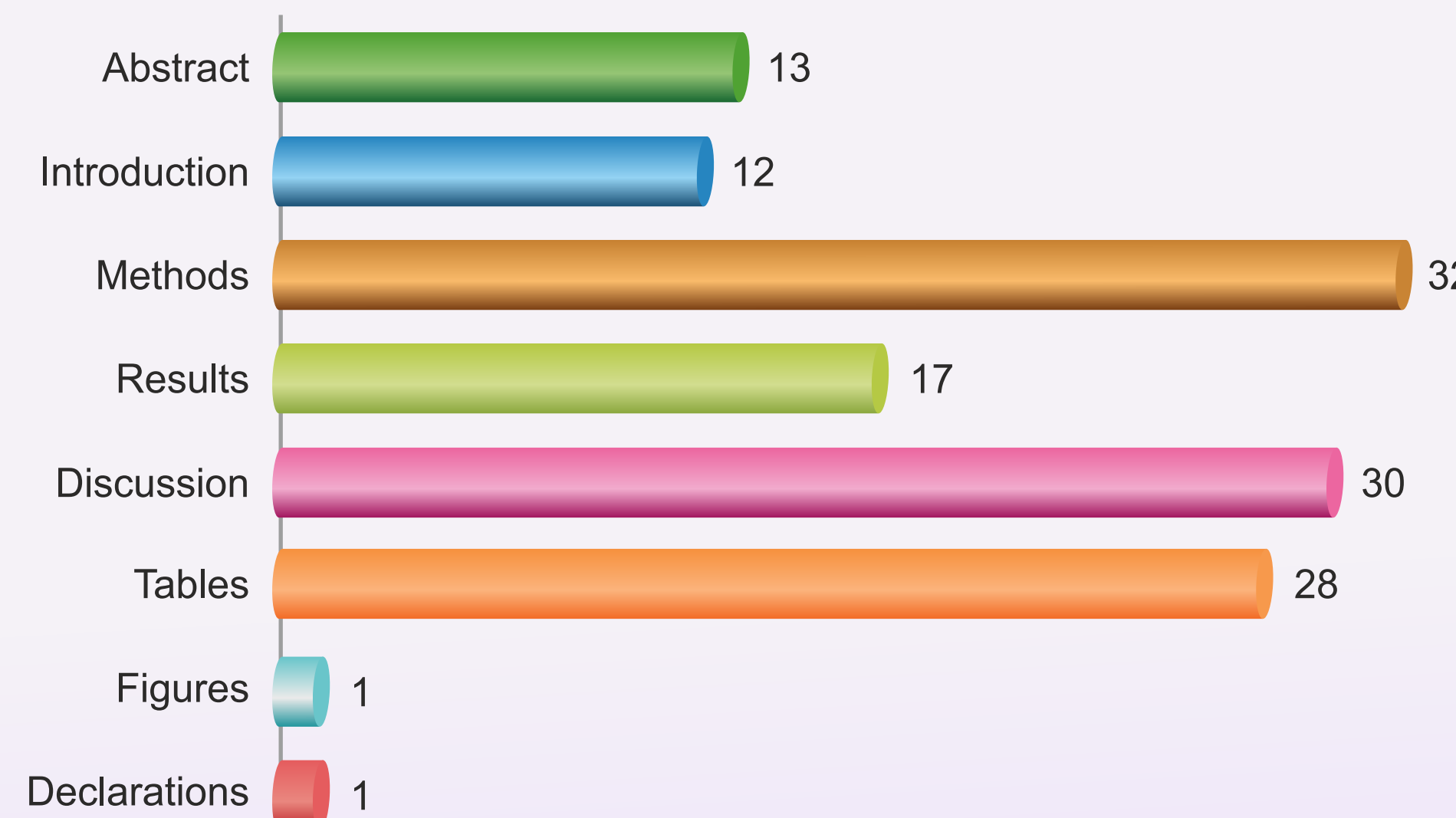
- Number of hits**
- Sources of published content**
- Nature of similar word-strings**
(common phrases/ definitions or verbatim copy/ paste)



RESULTS

The Similarity report (34%) captured hits (range, 1-32) for all sections of the manuscript even title, abstract, tables/figures and ethics statements/ acknowledgement text

Number of Similarity hits in each section of the manuscript



34% Similarity index included:

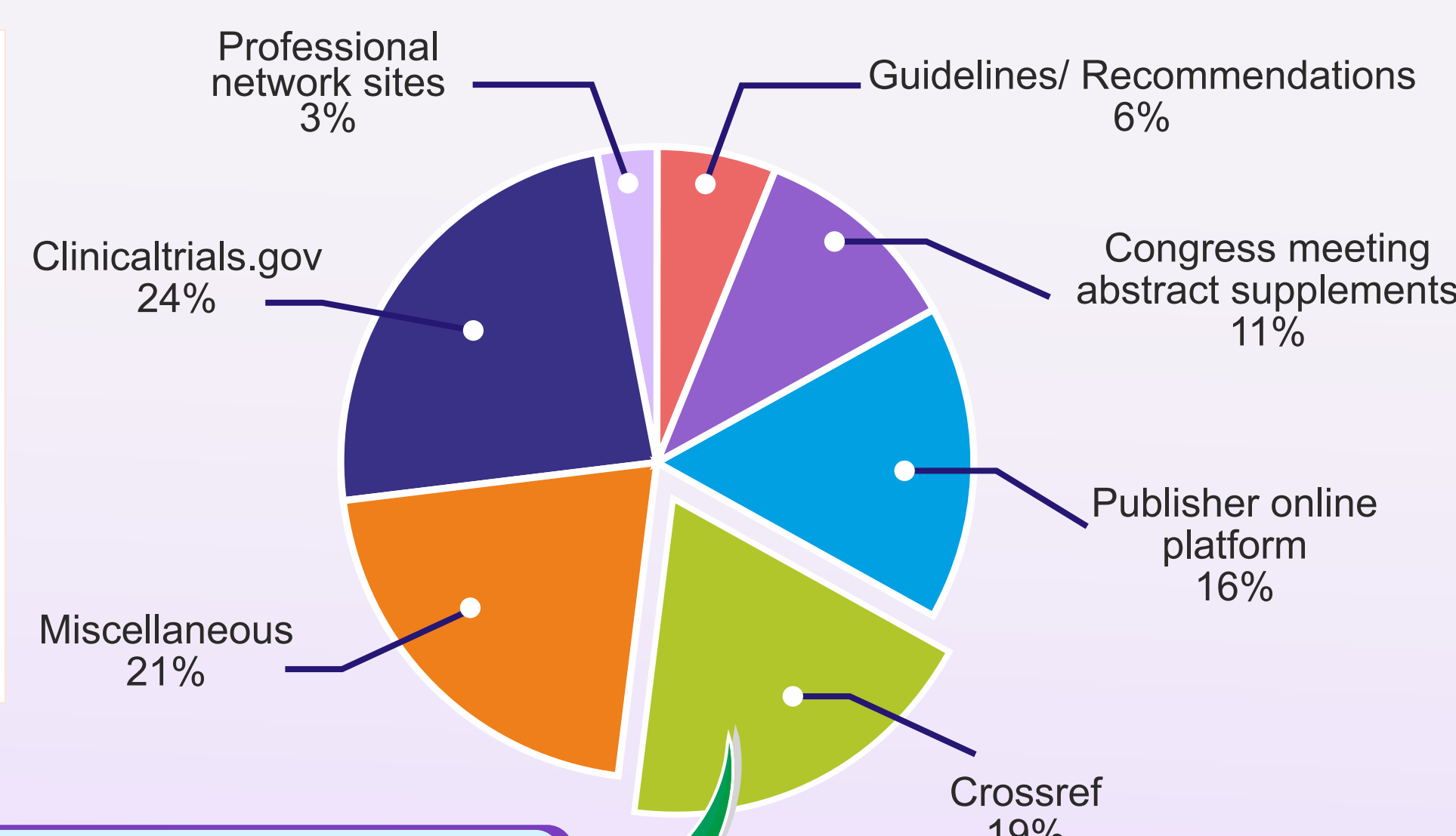
81%: From sites like Clinicaltrials.gov or guidelines or published congress abstract supplements

19%: From Cross-ref

*10%: Relevant articles

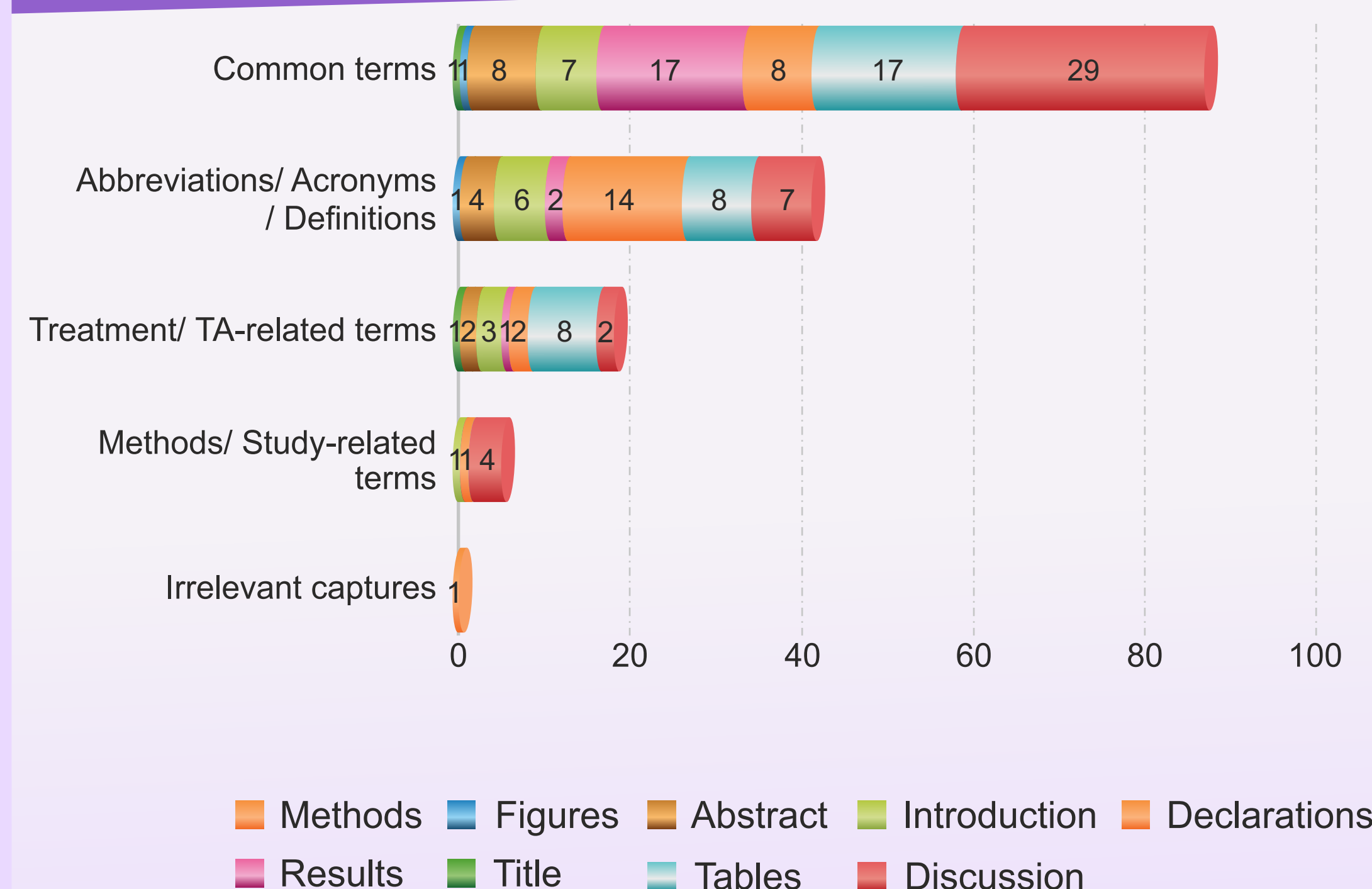
*9%: Irrelevant articles

34% SIMILARITY INDEX		
PRIMARY SOURCES		
1	www.clinicaltrials.gov	261 words — 6%
2	clinicaltrials.gov	97 words — 2%
3	"Abstract Supplement 2018 ACR/ARHP Annual Meeting", Arthritis & Rheumatology, 2018	78 words — 2%
4	link.springer.com	71 words — 2%
5	www.nice.org.uk	60 words — 1%
6	"2015 ACR/ARHP Annual Meeting Abstract Supplement", Arthritis & Rheumatology, 2015.	54 words — 1%



Relevant articles related to TA/ therapy: 10%
Irrelevant articles: 9%

Categories of Similarity



Abbreviations/ Acronyms/ Definitions:
Nonsteroidal anti-inflammatory drugs (NSAIDs)

Method/ study-related terms:
Open-label, real-world evidence, prospective, multicentre, non-comparative, interventional, phase

Treatment/ therapeutic area-related terms:
Week 24/ subcutaneous/ Ankylosing spondylitis

Common terms:
Safety and efficacy

Irrelevant captures: Acknowledgements/ funding/ disclosures standard text

Proposed Recommendations

- Analyze the report and discern plagiarized vs required repeated text
- Have a candid discussion with the stakeholders regarding the report and analysis
- Rephrase the text where possible without losing context
- Curate an appropriate response letter to the journal
 - Convey the reasons for unavoidable similarity and provide examples of over-sensitive detection
 - Provide some solutions to tackle the repetitive text in methods (shift to supplementary material)

Common terms captured for Similarity hits

In body text

In a real-world clinical setting
Safety and efficacy of XXX in YYY patient with active zzz disease
nonsteroidal anti-inflammatory drugs (NSAIDs)
The safety profile of XXX was consistent with the reported randomized-controlled global studies
This Phase 4 Single-arm, multicenter
American College of Rheumatology (ACR20)
None of the TEAEs led to discontinuation of the treatment
a higher proportion of patients achieved ASAS20 response at week 24 in this study
Treatment-emergent adverse events No deaths were reported
... included a screening phase, treatment phase, and follow-up phase
TEAEs were coded under the Medical Dictionary for Regulatory Activities (MedDRA), Version 24.0, was used
The study protocol XXX was approved by the institutional review boards ... ethics committees at the...the study
Pregnant or nursing women, ... women planning a pregnancy
aged ≥18 years with active

Acknowledgement:

provided writing assistance and
provided additional editorial support (both from SIRO Clinpharm Pvt. Ltd., India). The authors also thank the patients for their support.

In tables and figures

At week 24 -11.28 (20.69)
ysis set include the patients who were enrolled, received at least one dose week 24 visit. All values are expressed as mean (SD); Assessments were nging from 0 to 10, where 0 represents productivity not affected at all/no luctivity affected very much/the worst pain imaginable unless until specif College of Rheumatology 20% improvement criteria; AS, ankylosing s Spondyloarthritis International Society 20% improvement criteria; BASI ondylyitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Fi e protein; ESR, erythrocyte sedimentation rate; HAQ, Health Assessment Treat; NRS, numeric rating scale; PsA, psoriatic arthritis; SC, subcutaneo swollen joint count; TJC, tender joint count

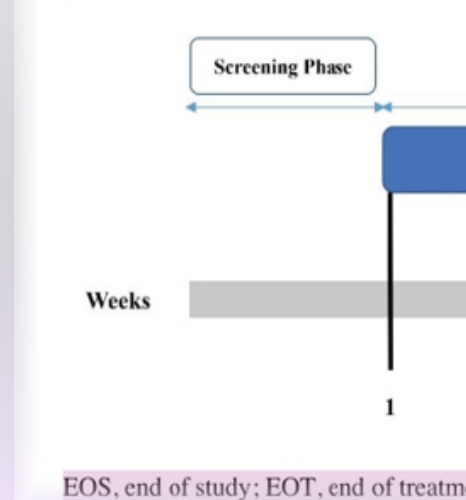
TEAEs leading to death

Serious TEAEs
Reasonably related serious TEAEs
TEAEs leading to discontinuation of study drug
TEAEs leading to termination of study participation
TEAEs of safety events of interest (special

SJC

Baseline score
Change from baseline
At week 14
At week 24
Patient's assessment of pain
Baseline score
Change from baseline
At week 14
At week 24
Patient's global assessment of disease activity
Baseline score
Change from baseline
At week 14

Figure 1: Study design



Acknowledgements

The authors would like to thank Ms. Mugha Rokade (SIRO Clinpharm Pvt Ltd.) for layout and graphics support.

Disclosures

All authors are employees of SIRO Clinpharm Pvt. Ltd.

Presented at the 20th Annual Meeting of ISMPP, April 29– May 01, 2024; Washington, DC