



**HAL**  
open science

## Design and methodological issues of within-person (split-body) randomized controlled trials evaluating a topical treatment: A systematic review

S Leducq, Amandine Dugard, Aude Allemang-Trivalle, Bruno Giraudeau, A Maruani

### ► To cite this version:

S Leducq, Amandine Dugard, Aude Allemang-Trivalle, Bruno Giraudeau, A Maruani. Design and methodological issues of within-person (split-body) randomized controlled trials evaluating a topical treatment: A systematic review. *Dermatology*, 2023, 10.1159/000530149 . hal-04106133

**HAL Id: hal-04106133**

**<https://hal.science/hal-04106133>**

Submitted on 25 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Research Article**  
***Design and methodological issues of within-person (split-body)***  
***randomized controlled trials evaluating a topical treatment: A***  
***systematic review***

Sophie Leducq<sup>a,b</sup>, Amandine Dugard<sup>a</sup>, Aude Allemang-Trivalle<sup>a</sup>, Bruno Giraudeau<sup>a,c\*</sup>, Annabel  
Maruani<sup>a,b\*</sup>

*\*equally contributed*

<sup>a</sup>University of Tours, Nantes University, INSERM, SPHERE 1246, Tours, France

<sup>b</sup>Department of Dermatology and Reference Center for Rare Diseases and Vascular Malformations (MAGEC), CHRU Tours, 37044 Tours Cedex 9, France

<sup>c</sup>INSERM CIC 1415, CHRU Tours, 37000 Tours, France

Short Title: Design and methodological issues of within-person trials

Corresponding Author:

Dr Sophie Leducq

University of Tours and Nantes University, INSERM, SPHERE 1246

2 boulevard Tonnellé

37000 Tours, France

Tel: +33 2 47 47 90 80 Fax: +33 2 47 47 82 47

Email: [sophie.leducq@univ-tours.fr](mailto:sophie.leducq@univ-tours.fr)

Key Message: Planification and reporting of within-person trials in dermatology are not optimal.

Number of Tables: 4

Number of Figures: 2

Word count: Please indicate the word count including Abstract and body text. This is not to include the title page, reference list or figure legends.

Keywords: topical treatment, within-person design, split body design, intra-individual comparison, randomized clinical trial

# 1 **Abstract**

## 2 *Background*

3 Topical drugs are often used as first-line treatment for dermatological conditions. A within-person  
4 design may then be well adapted: it consists of randomizing lesions/body sites rather than patients,  
5 which are then concomitantly treated by the different drugs compared, reducing inter-group  
6 variability and therefore requiring fewer patients than the classical parallel-group trial.

## 7 *Objectives*

8 The aim of this review was to provide a methodological overview of within-person randomized trials  
9 (WP-RCTs) in dermatology.

## 10 *Methods*

11 We searched for eligible trials published between 2017 and 2021 in MEDLINE, Embase and Central in  
12 dermatology journals and the 6 highest-impact-factor general medical journals. Two authors selected  
13 publications and extracted data independently.

## 14 *Results*

15 From 1034 articles identified, we included 54 WP-RCTs, mainly for acne vulgaris, psoriasis, actinic  
16 keratosis and atopic dermatitis. In most of the trials, patients had only 2 lesions/body sites. In none  
17 of the trials did we detect a potential carry-across effect (known to be the major methodological  
18 problem in WP-RCTs). Twelve studies reported a care provider applying the treatment, and in 26  
19 studies, the patients themselves applied the treatment. Finally, we also highlight statistical issues for  
20 the statistical analysis: overall, 14 (26.9%) studies used a test for independent observations, thus  
21 ignoring the between-lesion correlation.

## 22 *Conclusion*

23 Our systematic review highlights that despite the publication of the CONSORT checklist extension for  
24 WP-RCTs in 2017, this design is rarely used, and when it is, there are methodological and reporting  
25 concerns.

## 26 **Introduction**

27 Topical treatments are of major importance in dermatology and are the most prevalent drugs used  
28 for skin conditions.[1-4] They have an immediate local effect on skin lesions and are often used as  
29 first-line treatments to avoid systemic effects.[5,6] They can be administered alone or in combination  
30 with other topical or systemic treatments. Therapies with topical administration are of 3 subgroups:  
31 topical pharmacological drugs (e.g., topical corticosteroids, antibiotics, anti-inflammatory),  
32 emollients and cosmetics, and topical medical devices (e.g., bandages for leg ulcers). The choice of  
33 using a topical rather than a systemic treatment in skin conditions depends on several factors such as  
34 the severity of the disease and its extension.

35 Indeed, skin diseases might be classified according to the types of lesions and their extension to the  
36 body as 1) conditions that present as a single lesion (e.g., skin cancers); 2) conditions with multiple  
37 countable lesions (e.g., plaque psoriasis, eczema, vitiligo); and 3) multiple diffuse lesions extended to  
38 the body surface (e.g., scabies, viral exanthema). Thus, variability in the number and topography of  
39 skin lesions offers different methodological possibilities for randomized controlled trials (RCTs)  
40 assessing topical treatments.[7] The design of RCTs depends on the unit of randomization chosen  
41 (i.e., whether patients or lesions are randomized) but also whether a unique lesion or several lesions  
42 per patient are considered. According to these choices, RCTs of topical treatments might be designed  
43 as classical individual parallel RCTs,[8] individual parallel group with clustering RCTs (whereby  
44 patients, considered as clusters of lesions, are randomized, with several lesions per patient  
45 treated),[9] or within-person RCTs (WP-RCTs), also called split-body RCTs or the intra-individual  
46 comparison design.[10] In WP-RCTs, lesions rather than patients are randomized and different drugs  
47 are concomitantly administered to the same patient. This design is different from the cross-over  
48 design because distinct lesions are treated with distinct drugs administered at the same time rather  
49 than sequentially. By reducing inter-group variability (as in cross-over RCTs), the WP-RCT gains in  
50 efficiency and therefore requires fewer patients than the classical parallel-group trial.

51 However, the WP-RCT design has specific methodological issues. The main one is the risk of a carry-  
52 across effect, defined as the leakage of the treatment from one site (or lesion) to another that has  
53 another treatment or is not treated. Moreover, because lesions from an individual patient are more  
54 similar than lesions from different patients, there exists a correlation between assessment units (i.e.,  
55 lesions) that must be considered in WP-RCTs, both for sample size calculation and data analyses.  
56 Other specificities must be considered and are summarized in **Box 1**.[7] Therefore, WP-RCTs are

57 interesting designs for assessing topical drugs in skin conditions but are challenging to be set up;  
58 reporting guidelines (CONSORT extension) were elaborated for WP-RCTs.[11]

59 Here, by performing a systematic review, we aimed to assess whether trialists who conducted WP-  
60 RCTs considered all methodological and reporting specificities in order to highlight the points  
61 remaining to be improved.

62

## 63 **Materials and Methods**

64 We conducted a methodological systematic review and followed the standard Preferred Reporting  
65 Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (**Supplemental file 1**). The  
66 protocol was registered in the International Prospective Register of Systematic Review (PROSPERO  
67 record: CRD42022315516).

### 68 **Eligibility criteria**

69 We included only RCTs published in dermatology journals defined by the National Library of Medicine  
70 Catalog (journals referenced in the NCBI Databases) or in the 6 highest-impact-factor general medical  
71 journals. Congress abstracts and protocols were excluded. All RCTs evaluating a topical drug and  
72 using a within-person design for cutaneous diseases were included. We excluded studies of  
73 physiologic skin changes (e.g., melasma, scars, xerosis and folds), oral and ophthalmology  
74 pathologies, and anti-aging and cosmetic treatments. We included studies of topical treatments (e.g.,  
75 cream, gel, solution) with an active drug, emollient, herbal and plant local treatment and medical  
76 devices with an active substance but excluded those assessing physical treatment (e.g., laser,  
77 phototherapy).

### 78 **Information sources and search strategy**

79 We searched MEDLINE, Embase and Central from 01/07/2017 (date of publication of the CONSORT  
80 checklist extension for reporting WP-RCTs)[11] to 31/07/2021 for all RCTs evaluating a topical drug  
81 and using a within-person design in dermatology. Search algorithms were refined in collaboration  
82 with an expert methodologist (BG) and are presented in **Supplemental file 2**. We restricted the  
83 inclusion of articles to those published in English. The data on characteristics of patients, outcomes  
84 and adverse events were also collected from the results published at ClinicalTrials.gov  
85 ([www.ClinicalTrials.gov](http://www.ClinicalTrials.gov)) or any other register specified in the publication.

### 86 **Selection process**

87 Two reviewers (SL and AAT) selected the studies independently according to both the title and  
88 abstract and then the full text according to previously established inclusion criteria. In case of  
89 disagreement, consensus was determined by a discussion between the 2 reviewers or by consulting a  
90 third reviewer (BG/AM) if needed.

### 91 **Data extraction**

92 We designed a structured data extraction form using AirTable (**Supplemental file 3**) and extracted  
93 the following characteristics: name of the first author, publication year, journal, study recruitment  
94 characteristics (subcontinent and multicenter study), identification as a within-person design and  
95 term used by the authors, study design (objective, design justification by the authors, type of  
96 experimental and control treatment, type and level of outcome, blinding, prevention of carry-across  
97 effect, use of an intervention care provider), characteristics of patients (dermatologic condition, type  
98 of population, eligibility criteria for patients and lesions), statistical considerations (estimated sample  
99 size, statistical methods for analysis), report of results (number of patients and lesions included,  
100 patient flow chart, description of baseline characteristics, number of lost to follow-up, reporting of  
101 adverse events [adverse events, by body sites/areas or by patients]), ethical aspects (patient  
102 acceptability, methods to facilitate recruitment) and funding. Data were extracted by 2 reviewers (SL,  
103 AD) individually with blinding to each other's decisions. Disagreements between individual  
104 judgements were resolved by consulting 2 other reviewers (AM, BG) as for the study selection.

### 105 **Synthesis methods**

106 Data are summarized as number (percentage) for categorical data and median (interquartile range)  
107 for continuous data. Statistical analyses were performed with R 4.1.2.

## 108 **Results**

109 The results of this systematic review have been structured in line with the methodological points  
110 raised by Leducq et al.[7]

### 111 **Study selection**

112 The PRISMA flow diagram summarizing the screening process is in **Figure 1**. Of 1 034 references  
113 identified, 54 publications met the inclusion criteria.

114 **General characteristics of included studies (Tables 1 and 2)**

115 Most of the trials were single centre (n=31, 57.4%) and were superiority designed RCTs (n=50,  
116 92.6%). The most frequent dermatological conditions were acne (n=14, 25.9%), actinic keratosis  
117 (n=11, 20.4%), atopic dermatitis (n=8, 14.8%) and psoriasis (n=6, 11.1%). Most of the trials included  
118 adults only (n=45, 83.3%).

119 **General characteristics specifically related to the within-person design (Table 2)**

120 ***Identification and justification for using a WP-RCT***

121 Trials were generally identified as a WP-RCT (n=52, 96.3%). Identification of the trial as a WP-RCT was  
122 included in the title for 31 (57.4%) articles, in the abstract for 44 (88.0%, no abstract available for 5  
123 articles) and in the manuscript for 48 (88.9%). However, different terms were used to describe this  
124 design, and several could be used in the same article: split-body/split-face/split-side (n=31/52,  
125 59.6%), intra-individual/intra-subject/intra-patient (n=23/52, 44.2%), within-person/within-  
126 patient/within-subject (n=4/52, 7.7%), left to right/left vs right/right-left comparison/left right  
127 comparative (n=7/52, 13.5%), bilateral controlled clinical trial (n=1/52, 1.9%), per-patient (half face)  
128 (n=1/52, 1.9%), concurrent bilateral (n=1/52, 1.9%), side by side (n=1/52, 1.9%) and forearm  
129 controlled (n=1/52, 1.9%). Only 10 (18.5%) studies reported the rationale and justification for using a  
130 within-person design. Reasons evoked were to limit inter-individual variability (9 studies, 90.0%);  
131 reduce the number of patients to be included (5 studies, 50.0%); improve treatment adherence,  
132 without more explanation (one study); and because of the disease prevalence (rare skin disease for  
133 one study).

134 ***Methods for optimizing recruitment***

135 Five (9.3%) studies planned an extension period (i.e., a period when all lesions/body sites were  
136 treated with the experimental treatment after a first period during which the experimental and  
137 control treatment were assessed). Such an extension period is expected to facilitate the recruitment  
138 of patients as well as their adherence to the study protocol, which allows for explaining to them that  
139 eventually, all their lesions will be treated by the experimental treatment.

140 ***Ethical aspects***

141 One trial reported the patient's acceptability of the design and ethical aspects, and these concerned  
142 ethical and cosmetic issues because the patients received a placebo on one side.[12]



143 **Characteristics of the topical drugs (Tables 2 and 3)**

144 Most studies compared 2 treatments (n=48, 88.9%) and in 21 (38.9%) studies, the topical treatment  
145 assessed in the WP-RCT was an add-on treatment. The most frequent formulation of experimental  
146 treatment was cream (n=20, 37.0%). For the control group, an active drug was used in 23/53 (43.4%)  
147 studies, emollient cream in 12/53 (22.6%) and vehicle in 8/53 (15.1%). The most frequent regimen  
148 was 2 applications per day (n=18, 33.3%) for 15 to 29 days (n=17, 31.5%). In 21 (38.9%) studies, the  
149 topical treatment assessed was an add-on treatment. The standard treatment was a physical  
150 treatment in 14 (66.7%) studies, a topical treatment in 5 (23.8%) studies or a topical medical device,  
151 an oral treatment and a combination of a topical treatment and an oral treatment in 1 (4.8%) study  
152 each.

153 **Methodological characteristics of included studies (Tables 2 and 4)**

154 ***Eligibility criteria***

155 Eligibility criteria were specified for both patients and body sites/lesions in 18 (33.3%) studies and for  
156 patients only in 35 (64.8%). Most of the studies included 2 units (lesions) per patient (n=52, 96.3%).  
157 No trial featured a single lesion split into 2 areas.

158 ***Compliance***

159 Twelve (22.2%) studies involved a care provider in charge of applying the treatments, whereas for 26  
160 (48.1%) studies, treatments were applied by the patients themselves. Two (3.7%) studies reported  
161 monitoring of treatment application.

162 ***Blinding***

163 The following actors were blinded: patients in 25/46 trials (54.3%, unclear for 8 trials), care providers  
164 in 2/9 (22.2%, unclear for 3 trials) and outcome assessors in 39/45 (86.7%, unclear for 9 trials). For 23  
165 trials (53.5%, unclear for 11 trials), both patients and outcome assessors were blinded.

166 ***Detection of a potential carry-across effect***

167 None of the 54 included studies reported the detection of a potential carry-across effect.

168 ***Outcomes***

169 A primary outcome was clearly defined for 44 (81.5%) studies and was mostly a subjective outcome  
170 (n=32/44, 72.7%), such as physician global assessment or psoriasis area severity index, quality of life  
171 etc.[13]

## 172 ***Randomization***

173 Most studies randomized 2 different anatomic areas (n=41, 75.9%) or 2 distinct lesions (n=5, 9.3%).

## 174 ***Statistical methods***

175 A total of 33 (61.1%) articles described a sample size calculation. Six (11.1%) studies accounted for a  
176 within-patient between-lesions correlation, without reporting the value of the assuming correlation.

177 For trials with 2 units per patient (n=52), 33 (63.5%) performed a paired test, and one study fitted a  
178 mixed effects model to take into account between-lesion correlation. No study reported the within-  
179 patient between-lesions correlation coefficient estimate. Overall, 14 (26.9%) studies used a test for  
180 independent observations, thus ignoring the between-lesion correlation. For the 2 trials with > 2  
181 units (lesions) per patient, in one,[14] the statistical analysis did not take into account the within-  
182 lesion correlation and the other used only a descriptive statistical analysis.[15]

## 183 **Reporting characteristics of the included studies**

### 184 ***Patient flow and baseline characteristics***

185 Only 17 (31.5%) studies provided a patient flow chart with number of patients. Only 2 other (3.7%)  
186 studies reported lesions/body sites numbers associated with patient numbers. Overall, 34 (63.0%)  
187 articles reported lost to follow-up and dropout. Baseline data were described for patients in only  
188 48.2% (n=26) of the articles, for body sites/lesions in only one (1.9%) article and for both patients and  
189 body sites/lesions in 24 (44.4%) articles.

### 190 ***Harms***

191 Local adverse events were reported in 43 (79.6%) articles and general adverse events in 14 (25.9%);  
192 one study reported only general adverse events without reporting local adverse events. The median  
193 number of lesions/body sites with local adverse events was 5.0 (interquartile range, 3.0-9.5) and 3.0  
194 (interquartile range, 0.0-5.75) in the experimental and control group, respectively.

## 195 **Summary of methodological aspects to consider in WP-RCTs across all included studies (Fig. 2)**

196 In **Figure 2**, we synthesize the quality of the methodological items to consider (**Box 1**) for planning  
197 and reporting WP-RCTs presented as percentages across all included studies.

## 198 **Discussion**

199 In this methodological systematic review, we included 54 trials using a within-person design.

200 Dermatology conditions were mostly non-rare skin diseases in adults with multiple countable lesions  
201 with involvement of at least 2 different anatomical sites. Only about one third of the trials specified  
202 eligibility criteria for both patients and body sites/lesions. Most trials involved 2 lesions per patient.

203 Twelve studies used a care provider for applying the treatments (which may enhance compliance and  
204 limit group contamination), and in 26 studies, the treatment was applied by patients themselves.  
205 None of the trials assessed a potential carry-across effect.

206 No trial included a single lesion divided into 2 areas, although this type of design may be particularly  
207 interesting in rare skin diseases.[16] Indeed, because of the low prevalence of rare diseases, classical  
208 individual parallel RCTs are not adapted, and therefore, specific designs adapted to rare diseases  
209 might be considered, such as cross-over designs or stepped-wedge designs.[17] Also, the within-  
210 person design might be preferred to reduce the number of patients to be included. Similarly, few  
211 trials included > 2 lesions per patient even though this type of design allows for reducing the number  
212 of patients to be included by increasing the number of observations (lesions/body sites) per patient.

213 Regarding modalities of administration of treatment, because patients receive both treatments  
214 concomitantly, the use of an intervention care provider who applies the treatments, such a nurse,  
215 can limit if not avoid intergroup contamination. However, only a few trials used a care provider, and  
216 monitoring of treatment application was rarely reported. Moreover, the absence or negligible  
217 systemic passage of the experimental treatment is mandatory to avoid a carry-across effect. Avoiding  
218 a carry-across effect is a necessary prerequisite for using a within-person design. However, in our  
219 review, no study reported methods to detect a carry-across effect, such as plasma dosage of the  
220 drug, which, however, is not systematically available. Other factors should be considered for the risk  
221 of systemic passage of the experimental treatment such as the molecular weight of the experimental  
222 treatment,[18] the surface to be treated and the alteration of the skin barrier, which may increase  
223 dermal absorption and therefore systemic passage.[19] However, in the end, having so few studies  
224 reporting the use of a care provider and none testing for a potential carry-across effect questions  
225 whether trialists have well appreciated the risk of between-group contamination in a WP-RCT.

226 Local side effects were commonly described, as for any time a topical treatment is assessed, and  
227 were reported for 43 studies. In a WP-RCT, such side effects may compromise blinding as well as

228 compliance because each patient (who receives both treatments concomitantly) can directly  
229 compare the adverse events occurrence (but also efficacy) and therefore guess which lesion received  
230 the experimental/control treatment. In such a situation, an objective outcome is mandatory to limit a  
231 potential exaggeration of treatment effects.[20] Outcomes such as spectrophotometry or blinded  
232 evaluation by an independent outcome assessor using photographs in case of subjective outcomes is  
233 highly recommended.[13]

234 Our study highlights a lack of quality for reporting despite the publication of the CONSORT extension  
235 for WP-RCTs in 2017. This lack of reporting concerned common items related to RCTs (e.g., 39% of  
236 the studies did not report a sample size) but also specific items related to WP-RCTs. Several studies  
237 have assessed the quality of reports of RCTs and the findings are similar regardless of the study  
238 design and condition. For example, a systematic review performed in 2013 that described non-  
239 pharmacological interventions of RCTs found that of 137 interventions included, only 53 (39%) were  
240 adequately described.[21] More recently, in a study of interventional trials studying colorectal  
241 cancer, only 5/101 (5%) trials reported all selected CONSORT items completely.[22]

242 Only one article reported the ethical aspects of the within-person design, although it actually refers  
243 more to the equipoise principle than the within-person design. Indeed, in this study, the authors  
244 mentioned that using a placebo *“brings ethical and cosmetic issues, because the patients cannot be*  
245 *deprived of the treatment in one side”* and thus decided to use an active comparator.

246 Finally, we also highlight statistical issues for both sample size calculation and the statistical analysis.  
247 These findings were similar to previous publications of WP-RCTs in dentistry or ophthalmology.[23-  
248 26] When 2 lesions per patient were included, a classical pair-matched test was used, which is  
249 correct. Some trials (9, 16.7%) used mixed models, but this was to consider intra-lesion correlation  
250 between repeated assessments, not within-patient between-lesion correlation. Such an approach is  
251 incorrect.

252 This systematic review has some limitations. First, we limited the eligibility criteria to reports  
253 published only in English and between 2017 and 2021. Therefore, this review is not extensive but  
254 provides a recent description of published WP-RCTs. Second, the identification of trials as WP-RCTs  
255 was heterogenous and therefore, with our selection of key words, we might have missed some WP-  
256 RCTs. Finally, we did not contact the first authors to deal with missing data.

257 In conclusion, the WP-RCT is an interesting design and an alternative to the classical individual  
258 parallel RCT for investigating topical drugs. Although this design is well adapted for evaluating topical

259 treatment in skin conditions, only 54 trials were published between 2017 and 2021, with  
260 methodological and reporting concerns.

261

262 **Statements**

263 **Acknowledgement**

264 The authors thank Mrs Virginie Rochereau, University of Tours, for her technical assistance.

265 The authors are indebted to Laura Smales for English proofreading and manuscript editing.

266 **Statement of Ethics**

267 An ethics statement is not applicable because this study is based exclusively on published literature.

268 **Conflict of Interest Statement**

269 The authors declare no conflict of interest.

270 **Funding Sources**

271 None

272 **Author Contributions**

273 *Conceptualization:* SL, BG and AM

274 *Data Curation:* SL, AD and AAT

275 *Formal Analysis:* SL

276 *Methodology:* SL, BG and AM

277 *Software:* SL

278 *Supervision:* BG and AM

279 *Visualization:* SL

280 *Writing - Original Draft Preparation:* SL, BG and AM

281 *Writing - Review and Editing:* SL, AD, AAT, BG and AM

282 **Data Availability Statement**

283 The data extracted from included studies are available from the corresponding author on reasonable  
284 request.

## References

1. de Lusignan S, Alexander H, Broderick C, et al. Patterns and trends in eczema management in UK primary care (2009-2018): A population-based cohort study. *Clin Exp Allergy*. 2021 Mar;51(3):483-494.
2. Menter A, Korman NJ, Elmets CA, et al. Guidelines of care for the management of psoriasis and psoriatic arthritis. Section 3. Guidelines of care for the management and treatment of psoriasis with topical therapies. *J Am Acad Dermatol*. 2009 Apr;60(4):643-59.
3. Powell HB, Adamson AS. Medicare Part D payments for brand and generic drugs prescribed by dermatologists. *J Am Acad Dermatol*. 2018 Sep;79(3):575-577.
4. Song H, Adamson A, Mostaghimi A. Medicare Part D payments for topical steroids: Rising costs and potential savings. *JAMA Dermatol*. 2017 Aug;153(8):755-759.
5. Maruani A, Samimi M, Lorette G, Le Cleach L. Comparative effectiveness of topical drugs in dermatologic priority diseases: geometry of randomized trial networks. *J Invest Dermatol*. 2015 Jan;135(1):76-83.
6. Zhang M, Silverberg JI, Kaffenberger BH. Prescription patterns and costs of acne/rosacea medications in Medicare patients vary by prescriber specialty. *Am Acad Dermatol*. 2017 Sep;77(3):448-455.e2.
7. Leducq S, Caille A, Le Cleach L, et al. Research techniques made simple: randomized controlled trials for topical drugs in Dermatology: when and how should we use a within-person design? *J Invest Dermatol*. 2020 May;140(5):931-938.e1.
8. Ho N, Pope E, Weinstein M, Greenberg S, Webster C, Krafchik BR. A double-blind, randomized, placebo-controlled trial of topical tacrolimus 0.1% vs. clobetasol propionate 0.05% in childhood vitiligo. *Br J Dermatol*. 2011 Sep;165(3):626-32.
9. Cavalié M, Ezzedine K, Fontas E, et al. Maintenance therapy of adult vitiligo with 0.1% tacrolimus ointment: a randomized, double blind, placebo-controlled study. *J Invest Dermatol*. 2015 Apr;135(4):970-974.
10. Lepe V, Moncada B, Castanedo-Cazares JP, Torres-Alvarez MB, Ortiz CA, Torres-Rubalcava AB. A double-blind randomized trial of 0.1% tacrolimus vs 0.05% clobetasol for the treatment of childhood vitiligo. *Arch Dermatol*. 2003 May;139(5):581-5.
11. Pandis N, Chung B, Scherer RW, Elbourne D, Altman DG. CONSORT 2010 statement: extension checklist for reporting within person randomised trials. *BMJ*. 2017 Jun;357:j2835.
12. Saki N, Ahramiyanpour N, Heiran A, Alipour S, Parvizi MM. Efficacy of topical dimethyl sulfoxide (DMSO) 50% solution vs tretinoin 0.5% cream in treatment of patients with primary macular amyloidosis: A split-side single-blinded randomized clinical trial. *Dermatol Ther*. 2020 May;33(3):e13305.
13. Savović J, Jones H, Altman D, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess*. 2012 Sep;16(35):1-82.
14. Guttman-Yassky E, Ungar B, Malik K, et al. Molecular signatures order the potency of topically applied anti-inflammatory drugs in patients with atopic dermatitis. *J Allergy Clin Immunol*. 2017 Oct;140(4):1032-1042.e13.
15. Danilenko M, Stamp E, Stocken DD, et al. Targeting Tropomyosin Receptor Kinase in Cutaneous CYLD Defective Tumors With Pegcantratinib: The TRAC Randomized Clinical Trial. *JAMA Dermatol*. 2018 Aug;154(8):913-921.
16. Leducq S, Caille A, Barbarot S, et al. Topical sirolimus 0.1% for treating cutaneous microcystic lymphatic malformations in children and adults (TOPICAL): protocol for a multicenter phase 2, within-person, randomized, double-blind, vehicle-controlled clinical trial. *Trials*. 2019 Dec;20(1):739.

17. Cornu C, Kassai B, Fisch R, et al. Experimental designs for small randomised clinical trials: an algorithm for choice. *Orphanet J Rare Dis.* 2013 Mar;8:48.
18. Bos JD, Meinardi MM. The 500 Dalton rule for the skin penetration of chemical compounds and drugs. *Exp Dermatol.* 2000 Jun;9(3):165-9.
19. Kezic S, Nielsen JB. Absorption of chemicals through compromised skin. *Int Arch Occup Environ Health.* 2009 May;82(6):677-88.
20. Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol.* 2014 Aug;43(4):1272-83.
21. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ.* 2014 Mar;348:g1687.
22. Pellat A, Boutron I, Ravaud P. Assessment of transparency and selective reporting of interventional trials studying colorectal cancer. *BMC Cancer.* 2022 Mar;22(1):278.
23. Koletsis D, Fleming PS, Seehra J, Bagos PG, Pandis N. Are sample sizes clear and justified in RCTs published in dental journals? *PLoS One.* 2014 Jan;9(1):e85949.
24. Lee CF, Cheng AC, Fong DY. Eyes or subjects: are ophthalmic randomized controlled trials properly designed and analyzed? *Ophthalmology.* 2012 Apr;119(4):869-72.
25. Lesaffre E, Garcia Zattera MJ, Redmond C, Huber H, Needleman I; ISCB Subcommittee on Dentistry. Reported methodological quality of split-mouth studies. *J Clin Periodontol.* 2007 Sep;34(9):756-61.
26. Lesaffre E, Philstrom B, Needleman I, Worthington H. The design and analysis of split-mouth studies: what statisticians and clinicians should know. *Stat Med.* 2009 Dec;28(28):3470-82.



## Figure Legends

Fig. 1. Prisma flow chart

Fig. 2. Authors' judgements about each of the methodological items to consider for planning and reporting presented as percentages across all included studies.

## **SUPPLEMENTARY MATERIAL**

**Supplemental file 1.** PRISMA checklist

**Supplemental file 2.** Search strategy

**Supplemental file 3.** Extraction form of extraction and collection of data

**Supplemental file 4.** References of included studies