CrossMark

REVIEW

Evaluating clinical trial design: systematic review of randomized vehicle-controlled trials for determining efficacy of benzoyl peroxide topical therapy for acne

Sonia A. Lamel[1] · Raja K. Sivamani[2] · Maral Rahvar[3] · Howard I. Maibach[4]

Abstract Determined efficacies of benzoyl peroxide may be affected by study design, implementation, and vehicle effects. We sought to elucidate areas that may allow improvement in determining accurate treatment efficacies by determining rates of active treatment and vehicle responders in randomized controlled trials assessing the efficacy of topical benzoyl peroxide to treat acne. We conducted a systematic review of randomized vehicle-controlled trials evaluating the efficacy of topical benzoyl peroxide for the treatment of acne. We compared response rates of vehicle treatment arms versus those in benzoyl peroxide arms. Twelve trials met inclusion criteria with 2818 patients receiving benzoyl peroxide monotherapy treatment and 2004 receiving vehicle treatment. The average percent reduction in total number of acne lesions was 44.3 (SD = 9.2) and 27.8 (SD = 21.0) for the active and vehicle treatment groups, respectively. The average reduction in non-inflammatory lesions was 41.5 % (SD = 9.4) in the active treatment group and 27.0 % (SD = 20.9) in the vehicle group. The average percent decrease in inflammatory lesions was 52.1 (SD = 10.4) in the benzoyl peroxide group and 34.7 (SD = 22.7) in the

vehicle group. The average percentage of participants achieving success per designated study outcomes was 28.6 (SD = 17.3) and 15.2 (SD = 9.5) in the active treatment and vehicle groups, respectively. Patient responses in randomized controlled trials evaluating topical acne therapies may be affected by clinical trial design, implementation, the biologic effects of vehicles, and natural disease progression. "No treatment" groups may facilitate determination of accurate treatment efficacies.

Keywords Randomized controlled trial · Study design · Placebo effect · Placebo response · Acne · Topical administration · Vehicle delivery · Benzoyl peroxide

✉ Howard I. Maibach
maibachh@derm.ucsf.edu

[1] Department of Dermatology, New York Medical College, Metropolitan Hospital Center, New York, NY, USA

[2] Department of Dermatology, University of California, Davis School of Medicine, Sacramento, USA

[3] Department of Pathology, Loma Linda University, School of Medicine, Loma Linda, USA

[4] Department of Dermatology, University of California, San Francisco School of Medicine, 90 Medical Center Way Surge 110, Box 0989, San Francisco, CA 94143, USA

Introduction

Acne vulgaris is a multifactorial inflammatory dermatologic condition, comprising the most common presentation to dermatologist's offices and affecting over 85 % of individuals in their lifetime [1, 4, 5, 49, 50]. The primary site of pathology is the pilosebaceous unit, located in greatest preponderance in the skin of the face and torso [40]. Topical therapy is a mainstay in the treatment of acne providing direct action at the site of the pathology. Many topical medications are limited in their ability to bypass the hydrophobic barrier of the stratum corneum, requiring structural alterations and pairing with solvents and/or vehicles to enhance drug delivery and esthetic effects [11, 37].

Acne is often associated with significant negative impacts on quality of life [35]. This drives a continued need to improve the efficacy of available treatments, often resulting in emergence of new delivery systems. Many new vehicles not only improve drug delivery and efficacy, but

Springer

offer cosmetic enhancements and inherent biological activity that could improve the natural disease course of acne [11, 28]. In addition, efficacy investigations of acne therapies are commonly determined in clinical trials comparing responses of an active drug group to those in a "placebo-" or "vehicle"-treated group that operate under the assumption that all vehicles are equal. Furthermore, the transient nature of acne lesions and the varying severity of acne in each patient render efficacy estimations questionable.

Benzoyl peroxide (BPO) is one of the most frequently used acne treatments and after decades of its use, it retains its potent antimicrobial effect against *Propionibacterium acnes* without fostering resistant strains, even when used in combination with topical antibiotics [12–14, 22, 29]. While its exact mechanism is not entirely known, BPO is an oxidizing agent known for its antimicrobial and keratolytic activity. As with many topical medications, BPO must be formulated to enhance its penetration through the stratum corneum and permeation to its site of action, and because of its unique structure, vehicles can greatly affect its concentration and efficacy [18, 29]. Recent technological innovations have led to vehicles that deliver the active ingredient more efficiently and produce biological actions including the absorption of excess sebum from the skin surface [17, 21]. For instance, use of a microsphere-based vehicle allows for sustained release of BPO improving the tolerability and decreasing the concentration-dependent irritancy without compromising its efficacy, even at lower concentrations [38, 48]. Such augmentation of pharmaceuticals to treat acne means that "vehicle"-treated groups should not be synonymous with "placebo"-treated groups [19]. Because vehicle treatment groups have been shown to exhibit large effects in the topical treatment of acne and a true placebo should have no inherent activity against the disease process, at a minimum, "no treatment" groups may be needed as comparators in topical acne trials [8].

This article evaluates randomized controlled trials that use a "vehicle" control group to assess the efficacy of topical BPO in the treatment of acne. BPO is unique in that approval of its use for acne in the United States did not require proof of efficacy in double-blind randomized controlled trials. We hope to use the comparison of active drug and vehicle treatment groups to provide a framework to analyze acne clinical trial design and implementation to draw important conclusions as to how therapeutic efficacy measurements can be improved. In addition, we hope to elucidate how physicians can use this information and efficacy results from randomized controlled trials to better optimize the therapeutic regimens they recommend to patients.

Methods

Data sources and search strategy

We searched Embase, PubMed, Cochrane Library, and ClinicalTrials.gov for articles published or registered randomized controlled trials through April 11, 2012. The search terms utilized were "acne" and "acneiform eruptions" combined with "controlled trial", "placebo", and "randomized trial" combined with "benzoyl peroxide". All articles written in any language were evaluated and results were translated to English as needed.

Selection and outcomes

Potentially eligible trials identified through database searching were initially reviewed individually by title and available abstract by all authors independently and results were cross-checked for accuracy. Studies were excluded if the investigational treatment was not BPO, they were not clinical trials investigating efficacy, or acne was not the treatment indication. Trials were then screened based on study procedures, the quality of the clinical trial, and excluded if there was no "vehicle" or "placebo" treatment group, BPO was not an investigational drug, or efficacy was not objectively measured with lesion counts.

Data abstraction

Data were independently abstracted by two authors for our primary outcome measures for each trial and the results were compared to ensure accuracy. The primary outcomes of interest included the percentage of responders in both the active treatment group and the vehicle group after the initial, blinded, vehicle-controlled treatment phase. We classified the outcome measures, recorded the length of time of the treatment period, and recorded pertinent inclusion and exclusion criteria. The number of patients treated in each study arm was documented along with dropout rate, the percentage of BPO in the active formulation, vehicle characteristics, and schedule of administration. Data on study design were recorded.

Statistical analysis

All patients in the active BPO monotherapy and vehicle groups were pooled. The patients in each group achieving the outcome measures were added according to the intention-to-treat numbers and the weighted averages were tabulated. Outcomes for percent reduction in total lesion

count, inflammatory lesion count, and non-inflammatory lesion count were determined and compared between the active treatment and vehicle groups. The average randomization fraction of patients allocated to active treatment versus vehicle treatment groups for all studies was also determined. For trials where patients were also treated with combination formulations or other active treatment preparations, all patients were grouped into the active treatment group.

Results

Trial flow

There were 103 potentially relevant studies retrieved in our search of PubMed, Embase, and clinical trials.gov. Ninety-one trials were excluded. Of the trials excluded, 1 trial did not evaluate acne vulgaris, 24 were not randomized controlled trials, 21 did not include a BPO monotherapy arm, 37 did not use a placebo or vehicle group as a comparator, 1 did not use the same vehicle comparator that was in the BPO formulation, 2 did not evaluate the efficacy of BPO, and 5 were repeat studies. Trials with "no treatment" groups were not found. Twelve studies met all inclusion and exclusion criteria and were included in the study procedures.

Trial characteristics

In the 12 studies analyzed, 2818 patients were in benzoyl peroxide monotherapy treatment groups and 2004 were in vehicle treatment groups. Trial characteristics including the numbers of participants randomized to each study arm, the treatment regimen, the vehicle type, inclusion and exclusion criteria, and outcome measures are listed individually for each study in Table 1.

On average, the studies included patients between 12.9 and 31.1 years of age, with 9.7–106.3 non-inflammatory lesions, and between 14.6 and 60 inflammatory lesions. The average number of daily treatment applications was 1.5 and the average study duration was 10.4 weeks. The average randomization fraction of active treatment to vehicle groups was 4:1 (Table 2).

Outcomes

The weighted average percent reduction in total number of acne lesions was 44.3 (SD = 9.2) and 27.8 (SD = 21.0) for the active and vehicle treatment groups, respectively (Table 2). The average reduction in the mean number of non-inflammatory lesions was 41.5 % (SD = 9.4) in the active treatment group and 27.0 % (SD = 20.9) in the vehicle group. The percent decrease in inflammatory lesions was 52.1 (SD = 10.4) in the benzoyl peroxide group and 34.7 (SD = 22.7) in the vehicle group. The percent of participants achieving treatment success as designated by the study outcomes were 28.6 (SD = 17.3) and 15.2 (SD = 9.5) in the active treatment and vehicle groups, respectively (Fig. 1).

Discussion

The percentages of "vehicle" or "placebo" responders in RCTs evaluating the efficacy of BPO in the treatment of acne are remarkably high, especially when compared to those of the active treatment responders. In addition, the differences between the active drug and placebo group have continued to decrease, especially in more recently conducted trials. Furthermore, the standard deviation of the weighted outcome averages is large. Together, these factors indicate that the accuracy of determined efficacy values may be compromised.

The lack of clear disparity between the treatment and vehicle responders could be secondary to a variety of factors either inherent in clinical trials that contribute to a "placebo effect" or to a false treatment response. These factors act to blur the line between the active drug response, vehicle response, placebo response, and the response expected secondary to natural disease progression. Acne lesions are transient and the severity of an individual's acne changes over time. Because it is not realistic to monitor these changes more frequently or in the time periods when patients are not enrolled in trials, there is a need to decipher values for each group of responders to more accurately define efficacies of therapeutics [26].

The placebo effect is an important concept within clinical trials and efficacy determinations, and much research has been done to unravel the components of the placebo effect [2, 9]. Factors within topical acne RCTs potentially reinforcing placebo or vehicle responses include therapeutic ritual from frequent dosing and the direct application of medication or vehicle to the site of pathology, the physical characteristics of the medication, more frequent office visits during trial periods, and increased attention to skin care during study periods [20, 30]. In addition, patient expectations contribute to observed effects with any prior experiences with RCT involvement, ineffective treatments, and provider relationships potentially acting to alter patient responses [27, 32, 43, 44]. However, other factors may also be improving disease course of acne.

Interestingly, the fraction of participants randomized to vehicle treatment groups in acne trials is very low with the

Table 1 Trial characteristics

References	Treatment arms (no. intention-to-treat subjects/arm)	Vehicle	Regimen	Acne severity	Inclusion criteria	Exclusion criteria	Outcome measures
Eichenfield et al. [16]	CLNP 1.2 %–BPO 3 % (n = 327) CLNP 1.2 % (n = 328) BPO 3 % (n = 328) Vehicle (n = 332)	Gel	QD × 12 weeks	Investigator's Static Global Assessment 2 + (6-point scale, 0 = clear skin, 5 = highly inflammatory lesions predominate)	Age 12–45 17–60 IL 20–150 NIL <4 NCL	Pregnancy H/o IBD	Mean absolute reduction IL Mean absolute reduction NIL Mean absolute reduction total lesions 2 grade improvement in ISGA
Gollnick et al. [24]	Adapalene 0.1 %–BPO 2.5 % (n = 419) Adapalene 0.1 % (n = 418) BPO 2.5 % (n = 415) Vehicle (n = 418)	Gel	QD × 12 weeks	Investigator Global Assessment 3 (5-point scale, 0 = clear, 4 = severe)	Age 12+ 17–60 IL 20–150 NIL <2 NCL	Severe acne (req. systemic retinoids) Pregnancy	% patients achieving IGA 0/1 % decrease of IL % decrease of NIL % decrease of total lesions
Gold et al. [23]	Adapalene 0.1 %–BPO 2.5 % (n = 415) Adapalene 0.1 % (n = 420) BPO 2.5 % (n = 415) Vehicle (n = 418)	Gel	QD × 12 weeks	Investigator Global Assessment 3 (5-point scale, 0 = clear, 4 = severe)	Age 12+ 20–50 IL 30–100 NIL <2 NCL	2+ Nodulocystic lesions Acne conglobata Acne fulminans Secondary acne Severe acne (H/o systemic ABX, h/o retinoids) Pregnancy	% patients achieving IGA 0/1 % decrease of IL % decrease of NIL % decrease of total lesions
Thiboutot et al. [46]	CLNP 1.2 %–BPO 2.5 % (n = 797) CLNP 1.2 % (n = 812) BPO 2.5 % (n = 809) Vehicle (n = 395)	Gel	QD × 12 weeks	Evaluator Global Severity Score 3 or 4 (6-point scale, = clear, 5 = very severe)	Age 12+ 17–40 IL 20–100 NIL <3 NCL	Pregnancy Washout periods to control for prior medication use	% patients achieving EGSS 0/1 % patients achieving 2 grade EGSS improvement % decrease of IL % decrease of NIL % decrease of total lesions
Thiboutot et al. [45]	Adapalene 0.1 %–BPO 2.5 % (n = 149) Adapalene 0.1 % (n = 148) BPO 2.5 % (n = 149) Vehicle (n = 71)	Gel	QD × 12 weeks	Investigator Global Assessment 2–4 (6-point scale, 0 = clear, 5 = very severe)	Age 12+ 20–50 IL 30–100 NIL 0 NCL	Severe acne Pregnancy	% patients achieving IGA 0/1 % decrease of IL % decrease of NIL % decrease of total lesions
Leyden et al. [31]	CLNP 1 %–BPO 5 % (n = 120) CLNP 1 % (n = 120) BPO 5 % (n = 120) Vehicle (n = 120)	Gel	QD × 10 weeks	Pillsbury Classification System grade 2 or 3 (4-point scale, 1 = least severe, 4 = most severe)	Age 13–30 10–50 IL 10–100 NIL <4 NCL	H/o systemic ABX (within last 4 weeks) Pregnancy	Decrease in IL Decrease in NIL Decrease in total lesions Improvement in physician global evaluation
Tschen et al. [47]	CLNP 1 %–BPO 5 % (n = 95) CLNP 1 % (n = 49) BPO 5 % (n = 95) Vehicle (n = 48)	Gel	BID × 10 weeks	Pillsbury Classification System grade 2 or 3 (4-point scale, 1 = least severe, 4 = most severe)	Age 13–30 10–80 IL 10–100 NIL <4 NCL	H/o systemic ABX H/o systemic retinoids Pregnancy	Decrease in IL Decrease in NIL Improvement in physician global evaluation (11-point scale, 5 = clear, −5 = disease exacerbation)

Table 1 continued

References	Treatment arms (no. intention-to-treat subjects/arm)	Vehicle	Regimen	Acne severity	Inclusion criteria	Exclusion criteria	Outcome measures
Lookingbill et al. [33]	CLNP 1 %–BPO 5 % (n = 95) CLNP 1 % (n = 89) BPO 5 % (n = 92) Vehicle (n = 58)	Gel	QD × 11 weeks	None specified	Age 13–30 12+ IL 12+ NIL <4 NCL	Pregnancy Washout periods	Decrease in IL Decrease in NIL Improvement in physician global improvement score (5-point scale, 0 = worsening, 4 = excellent)
Hughes et al. [25]	Isotretinoin 0.05 % (n = 25) BPO 5 % (n = 26) Vehicle (n = 26)	Gel	BID × 8 weeks	Mild to moderate Acne grade (Leeds technique, 11-point scale, 0 = none, 10 = severe)	Age 14–29 15–100 IL 15–100 NIL <4 NCL	Pregnancy H/o anti-androgen contraceptives H/o systemic retinoids within 24 mos	Decrease in IL Decrease in NIL Change in acne grade (0–10)
Mills et al. [36]	BPO 10 % (n = 25) BPO 5 % (n = 27) BPO 2.5 % (n = 76) Vehicle (n = 25)	Gel	BID × 8 weeks	Mild to moderately severe	10+ IL	None specified	% decrease of IL Global assessment of improvement
Chalker et al. [7]	Erythromycin 3 %–BPO 5 % (n = 44) Erythromycin 3 % (n = 45) BPO 5 % (n = 44) Vehicle (n = 44)	Gel	BID × 10 weeks	Grade 3+ acne (cook)	Age 15–30 10+ IL Grade 3+ acne (cook)	Pregnancy Serious illness H/o estrogens within 3 mos H/o atopy Washout periods	% decrease of IL % decrease of comedones % decrease of papules % decrease of pustules
Ede [15]	Chlorhydroxyquinolone 0.25 %–BPO 5.5 %–hydrocortisone 0.5 % (n = 45) Chlorhydroxyquinolone 0.25 %–BPO 5.5 % (n = 54) BPO 5.5 % (n = 48) Vehicle (n = 49)	Lotion	2.5×/day × 4 weeks	Grade II or III	Age 14–24 Grade II or III acne	H/o systemic ABX H/o endocrine therapy	% decrease of total lesions % overall clinical improvement

The inclusion/exclusion criteria, treatment groups with the intention-to treat numbers, the vehicle type, treatment regimens, and the outcome measures are also listed individually for each study

Table 2 Results

| References	Baseline severity scores (%)	Percent success rate definition, # (95 % CI, *P)	Baseline average total lesion count (SD)	Percent reduction total lesion count (SD, *P)	Baseline average NIL count (SD)	Percent reduction NIL count (SD, *P)	Baseline average IL count (SD)	Percent reduction IL count (SD, *P)
Eichenfield et al. [16]	BPO ISGA 3 (65)	2 grade improvement in ISGA; BPO 40 (25.5–35.5, 0.016)	BPO 71.6 (29.8)	BPO 54.5 (34.9, 0.032)	BPO 44.6 (23.8)	BPO 49.8 (39.5, 0.102)	BPO 27.0 (10.4)	BPO 62.2 (42.5, 0.015)
	Vehicle ISGA 3 (61)	Vehicle 18 (13.7–21.9, <0.001)	Vehicle 70.5 (29.7)	Vehicle 39.4 (42.3, <0.001)	Vehicle 44.2 (24.4)	Vehicle 33.5 (48.9, <0.001)	Vehicle 26.3 (9.8)	Vehicle 49.8 (46.0, <0.001)
Gollnick et al. [24]	BPO IGA 3 (100)	Achieving IGA of 0/1; BPO 26.7 (-, <0.001)	BPO 74.0 (-)	BPO 48.2 (-, <0.001)	BPO 45.0 (-)	BPO 48.8 (-, <0.001)	BPO 26.0 (-)	BPO 61.9 (-, <0.001)
	Vehicle IGA 3 (100)	Vehicle 17.9 (-, <0.001)	Vehicle 76.0 (-)	Vehicle 37.1 (-, <0.001)	Vehicle 46.0 (-)	Vehicle 36.7 (-, <0.001)	Vehicle 26.0 (-)	Vehicle 45.5 (-, <0.001)
Gold et al. [23]	BPO IGA 3 (100)	BPO 22.2 (-, <0.006)	BPO 76 (-)	BPO 47 (-, <0.017)	BPO 46 (-)	BPO 44.1 (-, <0.017)	BPO 27 (-)	BPO 55.6 (-, <0.017)
	Vehicle IGA 3 (100)	Vehicle 11.3 (-, <0.006)	Vehicle 76 (-)	Vehicle 28 (-, <0.017)	Vehicle 46 (-)	Vehicle 29.5 (-, <0.017)	Vehicle 27 (-)	Vehicle 34.3 (-, <0.017)
Thiboutot et al. [46]	BPO GSS 3 (82.4); GSS 4 (17.6)	2 grade improvement in EGSS; BPO 26.0 (-, -)	BPO 72.6 (-)	BPO 41.6 (-, -)	BPO 46.8 (-)	BPO 37.4 (-, -)	BPO 25.8 (-)	BPO 47.5 (-, -)
	Vehicle GSS 3 (80.8); GSS (19.2)	Vehicle 16.5 (-, -)	Vehicle 70.2 (-)	Vehicle 26.2 (-, -)	Vehicle 44.0 (-)	Vehicle 24.0 (-, -)	Vehicle 26.1 (-)	Vehicle 29.0 (-, -)
Thiboutot et al. [45]	BPO IGA 2 (10.1); IGA 3 (85.2); IGA 4 (4.7)	Achieving IGA of 0/1; BPO 15.4 (-, 0.003)	BPO 74 (-)	BPO 35.6 (-, <0.001)	BPO 43 (-)	BPO 36.4 (-, <0.001)	BPO 28 (-)	BPO 43.6 (-, <0.001)
	Vehicle IGA 2 (18.3); IGA 3 (80.3); IGA 4 (1.4)	Vehicle 9.9 (-, 0.002)	Vehicle 78 (-)	Vehicle 31.0 (-, <0.001)	Vehicle 46 (-)	Vehicle 37.5 (-, <0.001)	Vehicle 29 (-)	Vehicle 37.8 (-, <0.001)
Leyden et al. [31]	BPO Grade II (83); Grade III (17)	Improvement in acne grade; BPO 31 (-, <0.009)	BPO 49 (28)	BPO 27.3 (-, -)	BPO 30 (19)	BPO 25.3 (-, -)	BPO 19 (9)	BPO 30.5 (-, -)
	Vehicle Grade II (87); Grade III (13)	Vehicle 10 (-, <0.009)	Vehicle 47 (23)	Vehicle 2.1 (-, -)	Vehicle 28 (15)	Vehicle 1.8 (-, -)	Vehicle 19 (8)	Vehicle -2.6 (-, -)

Table 2 continued

References		Baseline severity scores (%)	Percent success rate definition, # (95 % CI, *P)	Baseline average total lesion count (SD)	Percent reduction total lesion count (SD, *P)	Baseline average NIL count (SD)	Percent reduction NIL count (SD, *P)	Baseline average IL count (SD)	Percent reduction IL count (SD, *P)
Tschen et al. [47]			Mean global improvement score						
	BPO	–	2.7 (1.5, –)	–	–	–	47.1 (–, –)	25 (–)	49.2 (–, <0.034)
	Vehicle	–	2.1 (1.3, –)	–	–	–	34.1 (–, –)	25 (–)	42 (–, <0.022)
Lookingbill et al. [33]			% achieving global improvement score of good or excellent						
	BPO	–	41 (–, –)	89.9 (–)	32.0 (–, –)	59.2 (–)	30 (4, –)	30.7 (–)	39 (4, –)
	Vehicle	–	10 (–, –)	101.6 (–)	1.9 (–, –)	70.1 (–)	–11 (8, –)	31.5 (–)	5 (7, –)
Hughes et al. [25]			Mean % reduction in acne grade						
	BPO	0.79 (–)	51.9 (–, –)	91 (–)	50.1 (–, –)	48 (–)	50 (–, –)	43 (–)	51.2 (–, –)
	Vehicle	0.7 (–)	–1.7 (–, –)	92 (–)	–6.5 (–, –)	49 (–)	–4.1 (–, –)	43 (–)	–8.5 (–, –)
Mills et al. [36]									
	BPO	–	–	19.5 (–)	50.1 (–, –)	–	–	–	–
	Vehicle	–	–	13.7 (–)	17.6 (–, 0.01)	–	–	–	–
Chalker et al. [7]									
	BPO	–	–	–	–	–	30 (–, –)	–	56 (–, –)
	Vehicle	–	–	–	–	–	3 (–, –)	–	27 (–, –)
Ede [15]			% achieving improved overall clinical response						
	BPO	–	65 % (–, –)	10.5 (–)	37.1 (–, –)	–	–	–	–
	Vehicle	–	33 % (–, –)	10.4 (–)	–5.8 (–, –)	–	–	–	–

The success rates as designated by each study and the percent reduction in total, non-inflammatory, and inflammatory lesion counts for each treatment group are listed individually for each study

* P is the level of significance achieved, if available
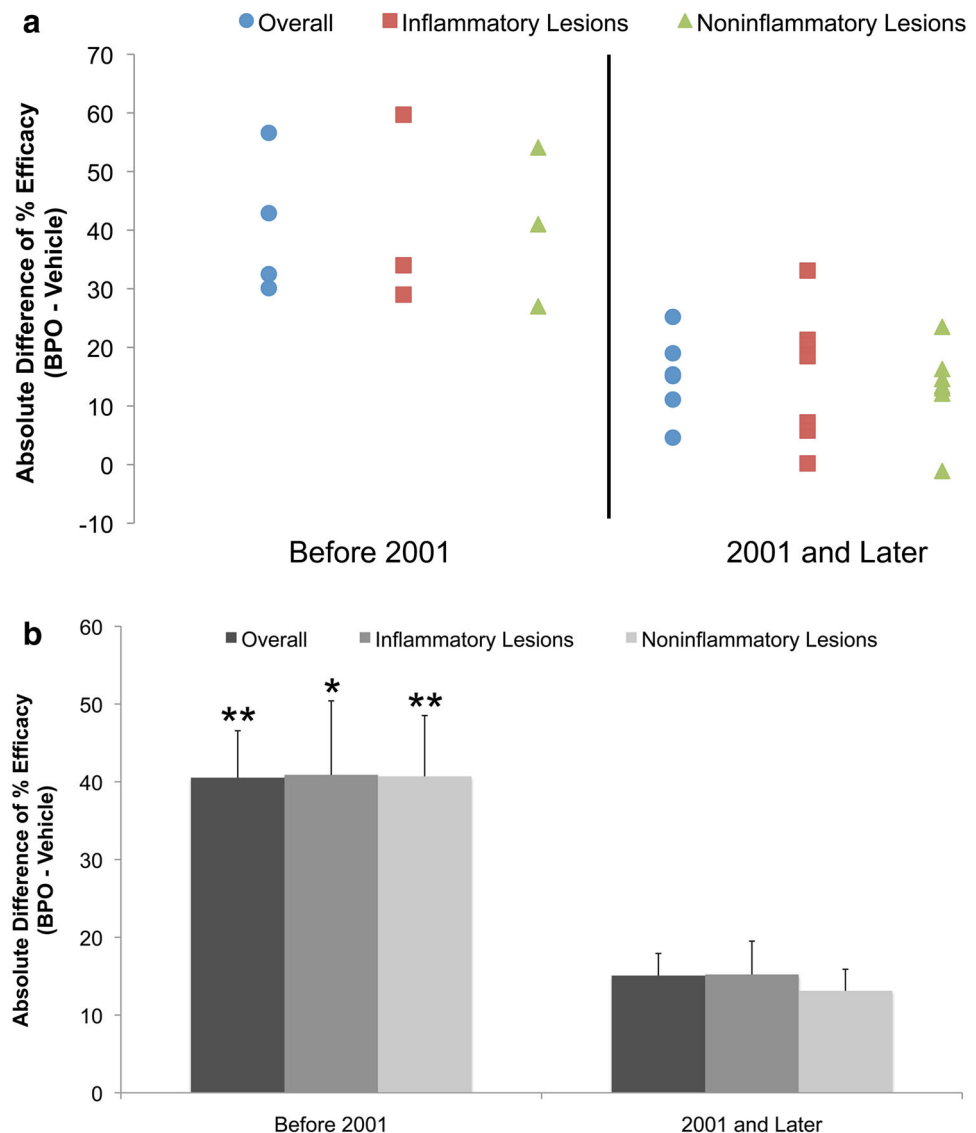
average ratio in this review being 1 to every 4 in the vehicle to active treatment group, respectively. It is important to note that patients are informed of the study design and the probability of receiving active drug prior to consent [3]. This leads to patient expectations of active drug receipt that may contribute to their response and skin care regimen adherence, and to placebo response. To more accurately assess efficacy, the placebo effect needs to be separated from vehicle effects. Moreover, many acne studies only evaluate treatment effects on the face, a site that is often visually inspected and critiqued, although potentially encouraging compliance with study protocols. This would escalate improvement in both treatment groups, further emphasizing the need for "no treatment" groups to have distinct values of drug efficacy, vehicle efficacy, placebo effect, and natural disease course.

BPO has been a mainstay of acne treatment for over five decades. Interestingly, only 12 published RCTs evaluated BPO monotherapy versus vehicle. Furthermore, only one of these trials compared vehicle to a BPO monotherapy formulation that is commercially available. The remaining 11 trials utilized BPO monotherapy formulated with a vehicle only available in combination products. With the wide variety of vehicles utilized in topical acne therapeutics and their biologic effects that potentially improve acne, the use of a "no treatment" group in RCTs is essential to determine true medication efficacies. Also, it is important to tease apart accurate efficacies of different drug monotherapies and their vehicles separately as treatment with combinations of monotherapies versus combination drug formulations can achieve equivalent results at a fraction of the cost [39].

There is great variability in the severity classifications used to determine patient eligibility in acne trials. In

addition, this variability also exists in the outcome measures captured. While the natural disease spectrum and course of acne is such that no consensus has been reached to homogenize acne severity classification, the design of RCTs determining medication efficacies should be standardized to allow for better comparison for different treatment efficacies [41, 42]. Furthermore, because outcome measures are determined by investigators, the accuracy of their assessments and determination of lesion counts are thus subject to discrepancy and error. Careful double blinding, scrutiny of observed treatment responses especially between the baseline and first follow-up visit, and the use of specialized imaging techniques to determine treatment responses could improve the accuracy of outcome measurements [34].

These findings must be interpreted in the context of our study design. Heterogeneity in RCT design and the differences in efficacy and tolerability of the BPO and/or vehicle formulations exist. Furthermore, there is yet to be a unified acne severity classification system or set of efficacy outcome measures. In addition, many efficacy trials evaluating acne treatments are comparator studies. We intend for the analysis of these trials to illustrate a variety of important points for consideration when interpreting efficacy results and designing of RCTs for topical acne treatments.

Importantly, our discoveries reveal the need for further investigation in multiple areas of RCT design and the separate implementation of vehicle and drug formulation efficacy testing. Furthermore, determination of the typical acne disease course is needed to validate the responses seen in RCTs. Standardization of acne severity measures for study inclusion and continued scrutiny of the accuracy of outcome measures can allow for cross comparison of different acne RCTs and improvement in our understanding of topical acne treatment efficacies. With continued testing of available acne treatments and a better understanding of responses, we can better guide patients.

References

1. Albuquerque RG, Rocha MA, Bagatin E, Tufik S, Andersen ML (2014) Could adult female acne be associated with modern life? Arch Dermatol Res 306:683–688

2. Beecher HK (1955) The powerful placebo. J Am Med Assoc 159:1602–1606

3. Benedetti F, Carlino E, Pollo A (2011) Hidden administration of drugs. Clin Pharmacol Ther 90:651–661

4. Bergfeld WF (2004) The pathophysiology of acne vulgaris in children and adolescents, part 2: tailoring treatment. Cutis 74:189–192

5. Borovaya A, Dombrowski Y, Zwicker S, Olisova O, Ruzicka T, Wolf R, Schauber J, Sardy M (2014) Isotretinoin therapy changes the expression of antimicrobial peptides in acne vulgaris. Arch Dermatol Res 306:689–700

6. Burkhart CG (2006) A chemical look at benzoyl peroxide may prove helpful. J Drugs Dermatol 5:705

7. Chalker DK, Shalita A, Smith JG Jr, Swann RW (1983) A double-blind study of the effectiveness of a 3 % erythromycin and 5 % benzoyl peroxide combination in the treatment of acne vulgaris. J Am Acad Dermatol 9:933–936

8. Chiou WL (2012) Low intrinsic drug activity and dominant vehicle (placebo) effect in the topical treatment of acne vulgaris. Int J Clin Pharmacol Ther 50:434–437

9. Colloca L, Miller FG (2011) How placebo responses are formed: a learning perspective. Philos Trans R Soc Lond B Biol Sci 366:1859–1869

10. Del Rosso JQ, Tanghetti E (2006) The clinical impact of vehicle technology using a patented formulation of benzoyl peroxide 5 %/clindamycin 1 % gel: comparative assessments of skin tolerability and evaluation of combination use with a topical retinoid. J Drugs Dermatol 5:160–164

11. Draelos ZD, Callender V, Young C, Dhawan SS (2008) The effect of vehicle formulation on acne medication tolerability. Cutis 82:281–284

12. Eady EA (1998) Bacterial resistance in acne. Dermatol 196:59–66

13. Eady EA, Farmery MR, Ross JI, Cove JH, Cunliffe WJ (1994) Effects of benzoyl peroxide and erythromycin alone and in combination against antibiotic-sensitive and resistant skin bacteria from acne patients. Br J Dermatol 131:331–336

14. Eady EA, Gloor M, Leyden JJ (2003) Propionibacterium acnes resistance: a worldwide problem. Dermatol 206:54–56

15. Ede M (1973) A double-blind, comparative study of benzoyl peroxide, benzoyl peroxide–chlorhydroxyquinoline, benzoyl peroxide–chlorhydroxyquinoline–hydrocortisone, and placebo lotions in acne. Curr Ther Res Clin Exp 15:624–629

16. Eichenfield LF, Alio Saenz AB (2011) Safety and efficacy of clindamycin phosphate 1.2 %-benzoyl peroxide 3 % fixed-dose combination gel for the treatment of acne vulgaris: a phase 3, multicenter, randomized, double-blind, active- and vehicle-controlled study. J Drugs Dermatol 10:1382–1396

17. Embil K, Nacht S (1996) The Microsponge Delivery System (MDS): a topical delivery system with reduced irritancy incorporating multiple triggering mechanisms for the release of actives. J Microencapsul 13:575–588

18. Fakhouri T, Yentzer BA, Feldman SR (2009) Advancement in benzoyl peroxide-based acne treatment: methods to increase both efficacy and tolerability. J Drugs Dermatol 657–661

19. Feldman SR, Fleischer AB Jr (2007) A method to gain a sense of relative efficacy of acne medications that have not yet been studied. J Drugs Dermatol 6:976

20. Feldman SR, Tan J, Poulin Y, Dirschka T, Kerrouche N, Manna V (2011) The efficacy of adapalene–benzoyl peroxide combination increases with number of acne lesions. J Amer Acad Dermatol 64:1085–1091

21. Fluhr JW, Barsom O, Gehring W, Gloor M (1999) Antibacterial efficacy of benzoyl peroxide in phospholipid liposomes. A vehicle-controlled, comparative study in patients with papulopustular acne. Dermatol 198:273–277

22. Ghali F, Kang S, Leyden J, Shalita AR, Thiboutot DM (2009) Changing the face of acne therapy. Cutis 83:4–15

23. Gold LS, Tan J, Cruz-Santana A, Papp K, Poulin Y, Schlessinger J, Gidner J, Liu Y, Graeber M (2009) A North American study of

adapalene–benzoyl peroxide combination gel in the treatment of acne. Cutis 84:110–116

24. Gollnick HP, Draelos Z, Glenn MJ, Rosoph LA, Kaszuba A, Cornelison R, Gore B, Liu Y, Graeber M (2009) Adapalene–benzoyl peroxide, a unique fixed-dose combination topical gel for the treatment of acne vulgaris: a transatlantic, randomized, double-blind, controlled study in 1670 patients. Br J Dermatol 161:1180–1189

25. Hughes BR, Norris JF, Cunliffe WJ (1992) A double-blind evaluation of topical isotretinoin 0.05 %, benzoyl peroxide gel 5 % and placebo in patients with acne. Clin Exp Dermatol 17:165–168

26. Ito S (2011) Placebo in clinical trials. Clin Pharmacol Ther 90:637–639

27. Kaptchuk TJ, Kelley JM, Conboy LA, Davis RB, Kerr CE, Jacobson EE, Kirsch I, Schyner RN, Nam BH, Nguyen LT, Park M, Rivers AL, McManus C, Kokkotou E, Drossman DA, Goldman P, Lembo AJ (2008) Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. Br Med J 336:999–1003

28. Kircik L, Friedman A (2010) Optimizing acne therapy with unique vehicles. J Drugs Dermatol 9:s53–s57

29. Kligman AM (1995) Acne vulgaris: tricks and treatments. Part II: the benzoyl peroxide saga. Cutis 56:260–261

30. Klinger R, Soost S, Flor H, Worm M (2007) Classical conditioning and expectancy in placebo hypoalgesia: a randomized controlled study in patients with atopic dermatitis and persons with healthy skin. Pain 128:31–39

31. Leyden JJ, Berger RS, Dunlap FE, Ellis CN, Connolly MA, Levy SF (2001) Comparison of the efficacy and safety of a combination topical gel formulation of benzoyl peroxide and clindamycin with benzoyl peroxide, clindamycin and vehicle gel in the treatments of acne vulgaris. Am J Clin Dermatol 2:33–39

32. Linde K, Witt CM, Streng A, Weidenhammer W, Wagenpfeil S, Brinkhaus B, Willich SN, Melchart D (2007) The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. Pain 128:264–271

33. Lookingbill DP, Chalker DK, Lindholm JS, Katz HI, Kempers SE, Huerter CJ, Swinehart JM, Schelling DJ, Klauda HC (1997) Treatment of acne with a combination clindamycin/benzoyl peroxide gel compared with clindamycin gel, benzoyl peroxide gel and vehicle gel: combined results of two double-blind investigations. J Am Acad Dermatol 37:590–595

34. Lucchina LC, Kollias N, Gillies R, Phillips SB, Muccini JA, Stiller MJ, Trancik RJ, Drake LA (1996) Fluorescence photography in the evaluation of acne. J Am Acad Dermatol 35:58–63

35. Mallon E, Newton JN, Klassen A, Stewart-Brown SL, Ryan TJ, Finlay AY (1999) The quality of life in acne: a comparison with general medical conditions using generic questionnaires. Br J Dermatol 140:672–676

36. Mills OH Jr, Kligman AM, Pochi P, Comite H (1986) Comparing 2.5 %, 5 %, and 10 % benzoyl peroxide on inflammatory acne vulgaris. Int J Dermatol 25:664–667

37. Nino M, Calabro G, Santoianni P (2010) Topical delivery of active principles: the field of dermatological research. Dermatol Online J 16:4

38. No authors listed (2009) Benzoyl peroxide microsphere formulations: what is the science supporting microsphere vehicle technology and clinical use? J Clin Aesthet Dermatol. 2(9):46–54

39. Ozolins M, Eady EA, Avery AJ, Cunliffe WJ, Po AL, O'Neill C, Simpson NB, Walters CE, Carnegie E, Lewis JB, Dada J, Haynes M, Williams K, Williams HC (2004) Comparison of five antimicrobial regimens for treatment of mild to moderate inflammatory facial acne vulgaris in the community: randomised controlled trial. Lancet 364:2188–2195

40. Pawin H, Beylot C, Chivot M, Faure M, Poli F, Revuz J, Dreno B (2004) Physiopathology of acne vulgaris: recent data, new understanding of the treatments. Eur J Dermatol 14:4–12

41. Pochi PE, Shalita AR, Strauss JS, Webster SB, Cunliffe WJ, Katz HI, Kligman AM, Leyden JJ, Lookingbill DP, Plewig G (1991) Report of the Consensus Conference on Acne Classification. Washington, D.C., March 24 and 25, 1990. J Am Acad Dermatol 24:495–500

42. Sanclemente G, Acosta JL, Tamayo ME, Bonfill X, Alonso-Coello P (2014) Clinical practice guidelines for treatment of acne vulgaris: a critical appraisal using the AGREE II instrument. Arch Dermatol Res 306:269–277

43. Stewart-Williams S, Podd J (2004) The placebo effect: dissolving the expectancy versus conditioning debate. Psycholog Bull 130:324–340

44. Stone DA, Kerr CE, Jacobson E, Conboy LA, Kaptchuk TJ (2005) Patient expectations in placebo-controlled randomized clinical trials. J Eval Clin Prac 11:77–84

45. Thiboutot DM, Weiss J, Bucko A, Eichenfield L, Jones T, Clark S, Liu Y, Graeber M, Kang S (2007) Adapalene–benzoyl peroxide, a fixed-dose combination for the treatment of acne vulgaris: results of a multicenter, randomized double-blind, controlled study. J Am Acad Dermatol 57:791–799

46. Thiboutot D, Zaenglein A, Weiss J, Webster G, Calvarese B, Chen D (2008) An aqueous gel fixed combination of clindamycin phosphate 1.2 % and benzoyl peroxide 2.5 % for the once-daily treatment of moderate to severe acne vulgaris: assessment of efficacy and safety in 2813 patients. J Am Acad Dermatol 59:792–800

47. Tschen EH, Katz HI, Jones TM, Monroe EW, Kraus SJ, Connolly MA, Levy SF (2001) A combination benzoyl peroxide and clindamycin topical gel compared with benzoyl peroxide, clindamycin phosphate, and vehicle in the treatment of acne vulgaris. Cutis 67:165–169

48. Wester RC, Patel R, Nacht S, Leyden J, Melendres J, Maibach H (1991) Controlled release of benzoyl peroxide from a porous microsphere polymeric system can reduce topical irritancy. J Am Acad Dermatol 24:720–726

49. White GM (1998) Recent findings in the epidemiologic evidence, classification, and subtypes of acne vulgaris. J Am Acad Dermatol 39:S34–S37

50. Younis S, Javed Q (2014) The interleukin-6 and interleukin-1A gene promoter polymorphism is associated with the pathogenesis of acne vulgaris. Arch Dermatol Res. doi:10.1007/s00403-014-1519-x: