Response to the Letter to the Editor Regarding "Crisaborole Ointment, 2%, for Treatment of Patients with Mild-to-Moderate Atopic Dermatitis: Systemic Literature Review and Network Meta-Analysis"

Kyle Fahrbach · Jialu Tarpey · Evelien Bergrath Washington ·

Rachel Hughes · Howard Thom · Maureen P. Neary · Amy Cha ·

Robert Gerber · Joseph C. Cappelleri

Dear Editors,

We thank you for the opportunity to respond to the concerns of Dr. Luger and Mr. Ranjan. Their concerns broadly fall into two categories: whether any key studies were missed, and whether the proper analytic methods were applied. We address each of these in turn.

First, regarding study inclusion, we respectfully disagree with the assessment that three studies were unduly excluded from analysis. As noted, the systematic review was designed to search for published studies investigating treatment for mild-to-moderate atopic dermatitis (AD) for which there were global assessments for Investigator's Global Assessment (IGA; performed at 4–6 weeks). For this reason, the unpublished Novartis study was not included. While both Leung et al. [1] and Emer et al. [2] studied patients with mild-to-moderate AD, Leung applied an additional, much more restrictive inclusion criterion: patients for

Evelien Bergrath Washington: Evidence Synthesis, Modeling and Communication, Evidera, 500 Totten Pond Road, Fifth Floor, Waltham, MA 02451, USA at the time of the original publication.

K. Fahrbach (✉) · J. Tarpey · E. B. Washington · R. Hughes
Evidence Synthesis, Modeling and Communication, Evidera, 500 Totten Pond Road, Fifth Floor, Waltham, MA 02451, USA
e-mail: kyle.fahrbach@evidera.com

J. Tarpey
e-mail: jialu.tarpey@evidera.com

E. B. Washington
e-mail: evelien.bergrathwashington@evidera.com

R. Hughes
e-mail: rachel.hughes@evidera.com

H. Thom
Bristol Medical School: Population Health Sciences, University of Bristol, 1-5 Whiteladies Road, Bristol BS8 1NU, UK

H. Thom
Clifton Insight, Bristol, UK
e-mail: howard.thom@bristol.ac.uk

M. P. Neary
Inflammation and Immunology, Pfizer Inc., 500 Arcola Road, Collegeville, PA 19426, USA
e-mail: Maureen.Neary@pfizer.com

whom topical corticosteroids (TCS) were clinically ineffective. It is likely that this fact led to their finding of an IGA response rate of 0 for vehicle, compared with the much higher rates seen in other studies. While disease was still considered mild-to-moderate, it is clear that this study was conducted on a notably different patient population than the other studies; inclusion of this study in our network meta-analysis (NMA) would have broken the similarity assumption that underlies NMA and thus would have led to biased results. Finally, Emer's study design was unique; both groups ($n = 20$ total) received both treatments, each applying them to opposite sides of the body, targeting one specific lesion on each side. The "Global Assessment" performed was therefore actually a local assessment and concerns only one lesion for each treatment. For this reason, the study was duly excluded from analysis. Additionally, note that neither Leung nor Emer were included in the Chia and Tey meta-analyses that the writers mention [3] and that Emer found essentially equal efficacy between pimecrolimus and the vehicle (72.5% versus 71.7%).

Second, regarding methods, Luger and Ranjan had four general concerns: the use of hazard ratios, a lack of correspondence to previous literature, the use of the baseline-risk model, and the possibility of alternate methods. We address each of these in turn.

To our knowledge, no study reported time-to-event data, so only an independent analysis of separate timepoints was feasible. We therefore chose the primary endpoints of all trials, with adjustment for different follow-ups, as our focus. It was these differences in follow-up that

A. Cha
Inflammation and Immunology, Pfizer Inc.,
235 East 42nd Street, New York, NY 10017, USA
e-mail: Amy.Cha@pfizer.com

R. Gerber · J. C. Cappelleri
Global Biometrics and Data Management
(Statistics), Pfizer Inc., 445 Eastern Point Road,
Groton, CT 06340, USA

R. Gerber
e-mail: robert.a.gerber@pfizer.com

J. C. Cappelleri
e-mail: joseph.c.cappelleri@pfizer.com

motivated our use of a complementary log–log model (cloglog) and hazard ratios; a naïve odds ratio analysis might have been biased because Investigator's Static Global Assessment (ISGA) success is not a rare event [4]. Use of hazard ratios is a direct consequence of using the cloglog model, which was done to help incorporate both 4-week and 6-week data; we make no claim as to whether that hazard ratio applies to earlier or later timepoints.

We believe that our results for pimecrolimus do not contradict those of the other published meta-analyses in any substantive way because we do find a very high likelihood that pimecrolimus is better than vehicle. The previously published meta-analyses cited by the letter writers are not focused specifically on the population with mild-to-moderate AD. The most recent meta-analysis, that of Chia and Tey (2015) [3], examines results regardless of baseline severity and thereby includes more studies with a generally different patient population; the same is true for Ashcroft et al. (2005) [5]. The other meta-analysis cited (Chen et al. 2010) [6] limited analysis to a pediatric population, whereas our analysis was in all patients ≥ 2 years of age.

The adjustment for baseline risk was not made solely because of expected differences in emollient. It was made because baseline risk inarguably varied strongly across trials owing to differences in a range of factors, including likely emollient ingredients, patient characteristics, and possible methodological factors. As we noted, some studies had higher rates for vehicle than were found in other studies for more active treatments, even though, within the study, active treatments always performed better than vehicle. There was strong evidence that the regression coefficient was nonzero [$b = -0.89$ (95% credible interval -1.26 to -0.47)] and, as we noted, the differences in deviance information criterion (DIC) were greater than 5, which are substantial based on the threshold of 5 suggested by general Bayesian analysis texts and of 3 by NMA-specific texts [7, 8]. Results for six models evaluated relative to baseline risk, class effects, and random/fixed effects are presented in the Supplementary Tables. Through observation of DIC and size of the slope, these

analyses show that only the three baseline risk models should be considered and that we chose our model as having the lowest DIC. Additionally, all of these baseline risk models (class effects and random effects) are very consistent in their conclusions.

Beyond the analyses for this NMA, it should also be noted that, overall, the need to adjust for baseline risk when considering relative response in autoimmune indications (of which AD is a part) has become increasingly apparent. It is quite relevant, for instance, that the exemplar baseline risk regression example in the National Institute for Health and Care Excellence guidelines is for an autoimmune condition (i.e., rheumatoid arthritis). We agree that the baseline adjustment provides an approximate correction, given the limited aggregated data available, but we believe that both from a clinical and a statistical standpoint this method of estimation is more accurate than that without the adjustment.

As a final note on methods, we appreciate the concern that, given the wide variability in vehicle rates, alternative approaches, such as an unanchored matched-adjusted indirect comparison (MAIC), should be considered. We presented a poster on this analysis at the European Academy of Allergy and Clinical Immunology Congress in July 2021, and we currently have an associated manuscript in preparation. The unanchored MAIC uses the same studies in mild-to-moderate AD as those studies included in our NMA. It found relatively strong evidence that crisaborole has higher odds of ISGA 0/1 than pimecrolimus 1% [odds ratio 2.03, 95% confidence interval (95% CI) 1.45–2.85, $p < 0.001$] and moderate evidence that it has higher odds than tacrolimus 0.03% (odds ratio 1.50, 95% CI 1.09–2.05, $p = 0.012$), while the tacrolimus 0.1% comparison was not feasible due to insufficient overlap in populations. We recognize, of course, that such an MAIC requires removing the vehicle arms from the analysis and, as a result, bias may arise if not all of the effect modifiers and prognostic factors are included in the adjustment.

We thank the authors for their comments on our analysis and agree with their conclusion that head-to-head clinical trials are essential.

We explicitly stated in our publication that "our results should be interpreted with caution and cannot replace a direct head-to-head evaluation." We think this qualification is especially true given the complex interplay between vehicle rates and relative effects versus vehicle, combined with the sparseness of the current network. Nevertheless, the results were based on the best available methods applied to the most applicable available evidence base.

ACKNOWLEDGMENTS

Roche Holding AG, Bayer AG, and Janssen Pharmaceuticals. Kyle Fahrbach, Jialu Tarpey, Evelien Bergrath, and Rachel Hughes are employed by Evidera, which provides consulting and other research services to pharmaceutical, medical device, and related organizations. In their salaried positions, they work with a variety of companies and organizations and are precluded from receiving payment or honoraria directly from these organizations for services rendered. Evidera received funding from Pfizer Inc. to conduct this study, develop the manuscript, and prepare this response letter.

Compliance with Ethics Guidelines. This letter is based on previously conducted studies and does not contain any studies with human participants or animals performed by any of the authors.

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

1. Leung DY, Hanifin JM, Pariser DM, et al. Effects of pimecrolimus cream 1% in the treatment of patients with atopic dermatitis who demonstrate a clinical insensitivity to topical corticosteroids: a randomized, multicentre vehicle-controlled trial. Br J Dermatol. 2009;161:435–43.

2. Emer JJ, Frankel A, Sohn A, Lebwohl M. A bilateral comparison study of pimecrolimus cream 1% and a topical medical device cream in the treatment of patients with atopic dermatitis. J Drugs Dermatol. 2011;10:735–43.

3. Chia BK, Tey HL. Systemic review on the efficacy, safety, and cost-effectiveness of topical calcineurin inhibitors in atopic dermatitis. Dermatitis. 2015;26: 122–32.

4. Thom H, Lopez-Lopez JA, Welton NJ. Shared parameter model for competing risks and different data summaries in meta-analysis: implications for common and rare outcomes. Res Synth Methods. 2019. https://doi.org/10.1002/jrsm.1371.

5. Ashcroft DM, Dimmock P, Garside R, Stein K, Williams HC. Efficacy and tolerability of topical pimecrolimus and tacrolimus in the treatment of atopic dermatitis: meta-analysis of randomized controlled trials. BMJ. 2005;330:516.

6. Chen SL, Yan J, Wang FS. Two topical calcineurin inhibitors for the treatment of atopic dermatitis in pediatric patients: a meta-analysis of randomized clinical trials. J Dermatol Treat. 2010;21:144–56.

7. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter DJ. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton, FL: CRC Press; 2012.

8. Dias S, Ades A, Welton N, Jansen JP, Sutton A. Network meta-analysis for decision-making. Hoboken, NJ: Wiley; 2018.