

Meta-analysis of clinical trials of probiotics for prevention and treatment of pediatric atopic dermatitis

Joohee Lee, AB,^b David Seto, BA,^b and Leonard Bielory, MD^a Newark, NJ

Background: Prenatal and postnatal probiotic supplementation for prevention and treatment of pediatric atopic dermatitis (PAD) has been studied in clinical trials, but results have been mixed and hindered by heterogeneity of study design.

Objectives: To summarize and interpret quantitatively clinical trial findings on the efficacy of probiotics for PAD and to define key trial features correlating with high methodologic quality.

Methods: PubMed and Cochrane database searches yielded 21 trials (n = 1898; age 0-13 y) published between February 1997 and May 2007 for review and quality assessment. Ten double-blind randomized controlled clinical trials were meta-analyzed by using RevMan. Data from the 6 prevention studies (n = 1581) and 4 treatment trials (n = 299) were pooled by using fixed-effects and random-effects models of relative risk ratios and of weighted mean difference, respectively.

Results: Prevention corresponded with summary effect sizes of 0.69 (0.57, 0.83) and 0.66 (0.49, 0.89), respectively, supporting probiotics' PAD prevention potential, which decreased further to 0.61 after exclusion of the 1 trial of postnatal-only probiotics. The clinical significance of the treatment trial findings of intergroup Scoring Atopic Dermatitis (quantification of PAD severity) score reduction by -6.64 points (-9.78, -3.49) and -8.56 (-18.39, 1.28), and intragroup change of -1.06 (-3.86, 1.73) and -1.37 (-4.81, 2.07), is questionable.

Conclusion: Current evidence is more convincing for probiotics' efficacy in prevention than treatment of PAD.

(J Allergy Clin Immunol 2008;121:116-21.)

Key words: Meta-analysis, pediatric atopic dermatitis, atopic dermatitis, probiotics, *Lactobacillus*, SCORAD, quality scales

Clinical investigation of probiotics' therapeutic potential in pediatric atopic dermatitis (PAD) reflects a growing interest in therapeutic targeting of events, namely intestinal colonization and T_H1/T_H2 immunologic maturation preceding the establishment of chronic atopic dermatitis (AD). Furthermore, AD is recognized as the earliest manifestation of atopy and therefore is an attractive target for efforts to halt this so-called *atopic march*.¹ Prenatal and/or postnatal probiotic supplementation for primary and secondary intervention for PAD has been studied in randomized

Abbreviations used

AD:	Atopic dermatitis
CR:	Clinical Relevance
ES:	Effect size
FE:	Fixed-effects
LGG:	<i>Lactobacillus rhamnosus</i> GG
MQ:	Methodological Quality
PAD:	Pediatric atopic dermatitis
RE:	Random-effects
RR:	Risk ratio
SCORAD:	Scoring Atopic Dermatitis
WMD:	Weighted mean difference

clinical trials, but results have been mixed. Assessing probiotics' clinical efficacy in preventing or treating PAD is hindered by subtle but potentially significant heterogeneity of protocols, variety of studied strains, and small sample sizes.

Numerous systematic reviews of probiotics for PAD have been published.²⁻⁵ Most of these publications have concluded that more work has to be done in this area before any clinical recommendations can be made. Our meta-analysis strives to take an additional analytic step beyond the qualitative assessments of systematic reviews. Although the various permutations of environment, host, probiotic strain, disease, and study design are obstacles for a robust quantitative synthesis, it is the premise of this meta-analysis to take a preliminary step in quantifying the current evidence. Albeit premature, this serves an important purpose. Many of the studies on probiotics and PAD are small trials. Meta-analysis is useful because it has the potential to generate hypotheses and assist in the design of large trials.⁶

This meta-analysis used fixed-effects (FE) and random-effects (RE) models, both of which fall under ANOVA, the standard statistical toolbox for meta-analyses. The merits and drawbacks of the FE and RE approaches to meta-analyzing data have been extensively discussed, especially in the field of social sciences. The differences between these 2 models lie in their assumptions about the process that generated the effect size (ES). They also differ in data analysis and in the interpretation of the results. Of interest to the clinician, the FE method is predisposed to type I bias in significance tests for mean ES. Although FE analyses are considered to be easier to understand and interpret, the RE methods tend to be more realistic in their assumptions and are less likely to overstate observed differences of interventions.⁷ Given these factors, combined with the nature of this meta-analysis, which we acknowledge as being potentially premature, we have chosen to rely on the RE model as the basis for our conclusions.

The objectives of this meta-analysis are as follows:

1. To summarize clinical trial findings on probiotics for PAD.

From ^athe Department of Medicine, ^bUniversity of Medicine and Dentistry of New Jersey-New Jersey Medical School.

Disclosure of potential conflict of interest: The authors have declared that they have no conflict of interest.

Received for publication August 6, 2007; revised October 2, 2007; accepted for publication October 26, 2007.

Reprint requests: Leonard Bielory, MD, Department of Medicine, UMDNJ-New Jersey Medical School, 90 Bergen Street, DOC Suite 4700, Newark, NJ 07103-2499. E-mail: bielory@umdnj.edu.

0091-6749/\$34.00

© 2008 American Academy of Allergy, Asthma & Immunology
doi:10.1016/j.jaci.2007.10.043

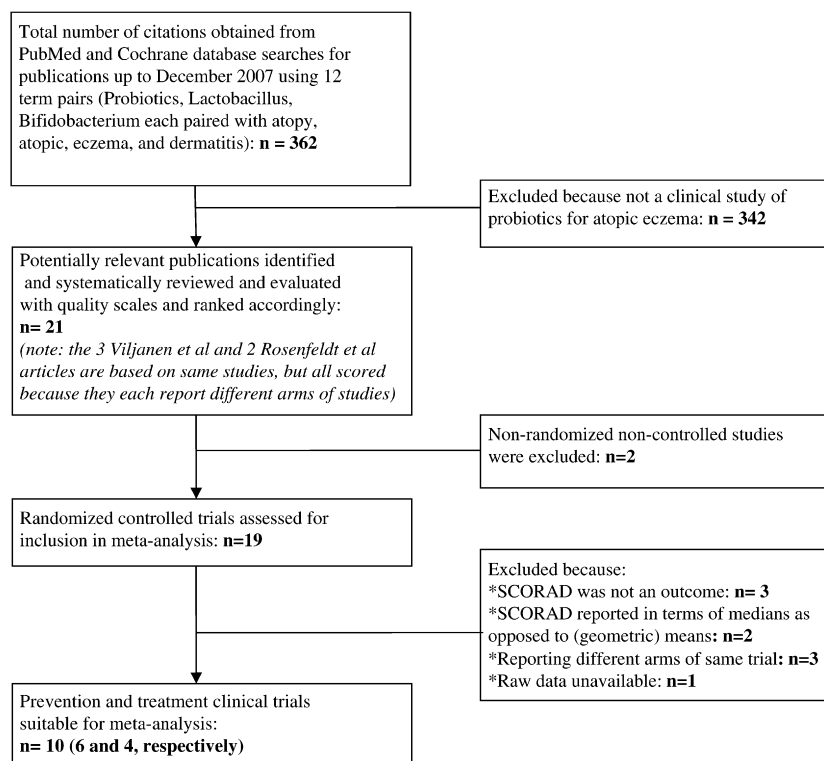


FIG 1. Flowchart of process of study selection for meta-analysis.

2. To apply 4 quality scales in scoring the clinical trials to distill features of study design correlating with high methodologic quality (see this article's Table E1 in the Online Repository at www.jacionline.org).⁸⁻¹⁰
3. To integrate quantitatively the available evidence on the value of probiotics in PAD prevention and treatment.

METHODS

Search strategy

PubMed and Cochrane databases were searched by using the following 12 pairs of search terms: *probiotics and atopy/atopic/eczema/dermatitis*; *Lactobacillus and atopy/atopic/eczema/dermatitis*; and *Bifidobacterium and atopy/atopic/eczema/dermatitis*. References from review articles and clinical trials were cross-checked to ascertain all published clinical trials were included in the scoring. Unpublished trials and abstracts were not included because scoring would not be possible. The most recent search was conducted in July 2007.

References from retrieved publications were cross-checked for additional potentially eligible studies in the scoring. Unpublished trials and abstracts were not included. Three of 5 principal investigators contacted for additional data provided the requested information not offered in the original publications.

Trial protocol features and significant findings of the studies were identified and presented in table format (see this article's Table E2 in the Online Repository at www.jacionline.org). The studies were evaluated by 2 blind reviewers (J.L. and D.S.) and refereed by a third party (L.B.) using the Cho & Bero Methodological Quality (MQ) and Clinical Relevance (CR) scales (Table E1). Each study fell into 1 of 4 categories: clinical evaluation, laboratory marker evaluation, both, or other. Two studies were discarded because one was observational and the other was questionnaire-based. The remaining studies were

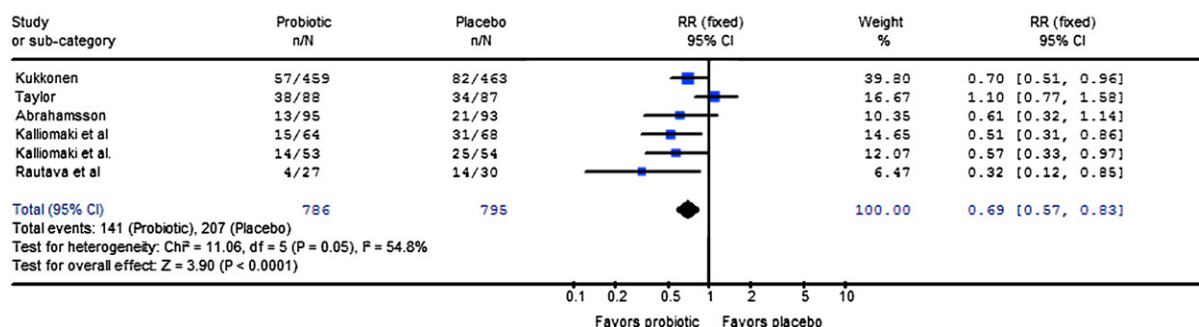
scored and ranked according to the Jadad and Delphi scales by each reviewer. Scores were used to rank the studies based on each of the scales. The ranked lists were analyzed for concordance between reviewers and across scales.

All double-blind randomized controlled trials of probiotics for PAD prevention or treatment were considered for meta-analysis (Fig 1). Outcomes from the prevention studies were presented in terms of postinterventional incidence of PAD in placebo and probiotics groups. Relative risk ratios (RRs) and their respective 95% CIs were taken as reported in the publications (Fig 2). Treatment trials assessed outcome in terms of reduction of PAD severity as quantified by the Scoring Atopic Dermatitis (SCORAD) system (SCORAD scores range from 0 to 103; higher scores correlate with greater severity). Ten of the 13 treatment studies defined clinical response, as quantified by reduction in SCORAD score, as an endpoint. Of these, only 4 were able to be meta-analyzed because there were publications of different arms of the same 2 trials ($n = 3$), SCORAD was reported in terms of medians as opposed to means ($n = 2$), and the necessary raw data could not be obtained ($n = 1$). These 4 studies used SCORAD scores to monitor trends within probiotics and placebo groups over the time of intervention (ie, intragroup change) and for head-on comparisons of disease severity between treatment groups after intervention (ie, intergroup difference). This meta-analysis pooled results with respect to both of these scopes of clinical response as published and/or as obtained from individual principal investigators (Fig 2).

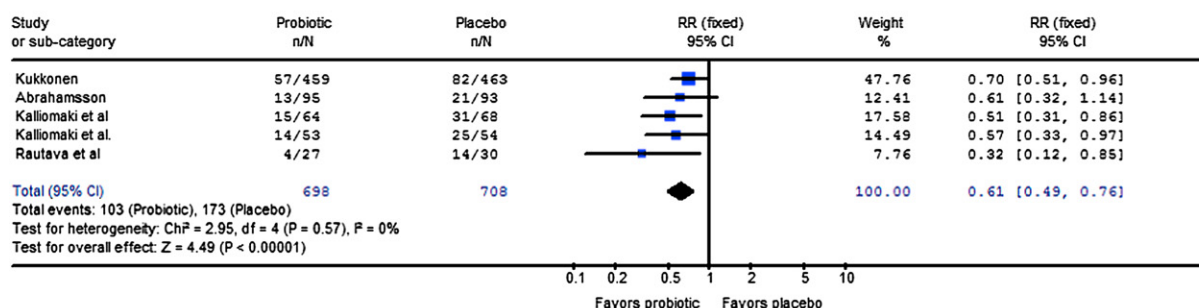
RESULTS

Twenty-one clinical studies were systematically reviewed (Table E2) and then scored and ranked by using 4 quality scales that ranged from 5 to 24 questions (Table E1). The studies varied in design but were predominantly randomized control trials. The 19 randomized double-blind clinical trials were subdivided on the basis of whether they were using probiotics for PAD prevention ($n = 6$) or treatment ($n = 13$). Each study defined its respective

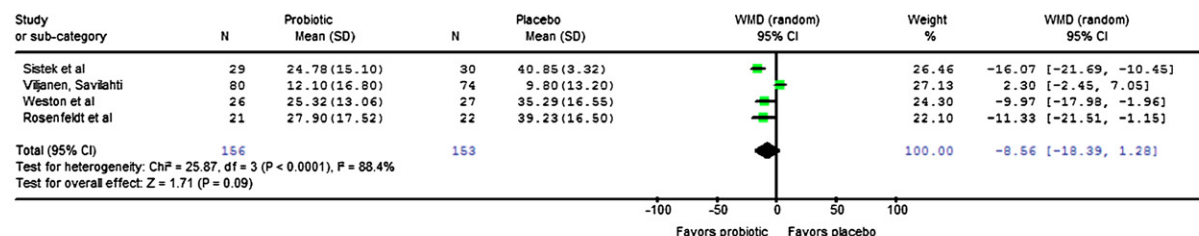
Review: Probiotics for pediatric atopic dermatitis (Version 4)
Comparison: 01 Probiotic versus placebo
Outcome: 04 Primary prevention studies



Review: Probiotics for pediatric atopic dermatitis (Version 4)
Comparison: 01 Probiotic versus placebo
Outcome: 03 Prenatal and postnatal prevention



Review: Probiotics for pediatric atopic dermatitis (Version 4)
Comparison: 01 Probiotic versus placebo
Outcome: 01 Final SCORAD



Review: Probiotics for pediatric atopic dermatitis (Version 4)
Comparison: 01 Probiotic versus placebo
Outcome: 02 SCORAD change score

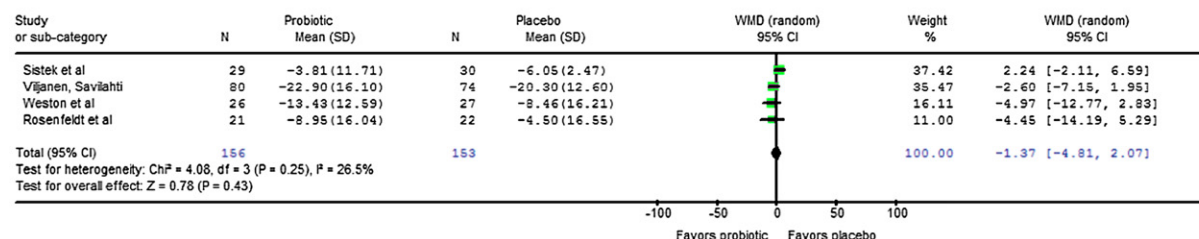


FIG 2. Pooled statistical analyses of primary and secondary prevention trials using RevMan. Prevention studies were assessed by using both FE and RE models of relative risk. The treatment studies were assessed in terms of intragroup and intergroup SCORAD change also using both models of WMD. Studies are ordered from top to bottom in descending order of MQ quality ranking.

inclusion and/or exclusion criteria with varying degrees of stringency.

A total of 1898 subjects, ages ranging from newborn to 13 years, were subjected to placebo or probiotics supplementation in an identical pill or powder in the 19 randomized clinical trials with prevention and treatment designs. The majority of trials

($n = 10$) chose *Lactobacillus rhamnosus* GG (LGG) as their probiotic strain under investigation. Five of the 6 prevention trials detected significant reduction of postintervention PAD incidence at 2 years of age. Five of the 10 treatment trials detected significant intergroup and/or intragroup SCORAD reductions, including 2 by subgroup analysis in terms of IgE sensitization.¹¹⁻¹⁴

Interestingly, drops in SCORAD scores have been observed to be a result primarily of a reduction of the extent, versus intensity or subjectivity, of disease.¹³⁻¹⁵

Results of quality scale application

The second reviewer (D.S.) consistently awarded studies lower absolute scores on the MQ than the first reviewer (J.L.) did, but there was a high degree of concordance between reviewers in the relative order of the rankings (Fig 3). There was no such obvious concordance between reviewers for the CR rankings. There were abundant ties in the scores for the CR and Jadad scales.

Results from the application of the 4 quality scales were analyzed for interobserver and interscale agreement (see this article's Fig E1 in the Online Repository at www.jacionline.org). The MQ came out as the clear victor in terms of interobserver reliability at $\kappa = 0.84$ (a figure that increased further to 0.875 with 5×5 κ tables in quintiles). In summary, there is statistical evidence that the longer MQ is very effective, and that the Jadad and Delphi may still be considered serviceable tools if researchers do not have the time or effort to spend on the MQ. The CR does not appear very useful as an inclusion tool.

The distribution of prevention versus treatment trials according to their quartile placements and raw scores on the Cho & Bero MQ scale illustrates the relative reporting quality of these 2 study types (Fig 3). Among the 6 prevention, 13 treatment, and 2 miscellaneous studies, the prevention studies unequivocally appear to be of the highest quality (Fig 3). A majority of the studies ranked by both reviewers to be in the first quartile on the MQ were prevention studies (3 of 5 studies). Conversely, a majority of those placed in the second and third quartiles were treatment studies (3 of 5 and 4 of 5 studies, respectively). None of the prevention studies fell in the fourth quartile; all of those that ranked in this bottom quartile were either treatment or miscellaneous studies. In addition, the mean raw MQ score of the primary prevention studies was 0.909, whereas the average raw score of the secondary studies was 0.808 ($P = .0004$ on unpaired t test, 95% CI), strongly suggesting that prevention studies in this meta-analysis routinely exemplified superior reporting as rated by the MQ scale.

Meta-analysis was done on 10 of the 19 randomized controlled trials primarily on the basis of data amenability to statistical pooling (ie, relative RR and [geometric] mean SCORAD scores). Reasons for exclusion included reporting of different arms of same trial (3 were eliminated because they were reports of serologic and fecal outcomes from 2 trials, which used SCORAD as their primary outcome) and data reporting deemed as inappropriate for statistical pooling, such as in terms of medians as opposed to (geometric) means. Although medians can be pooled, the users' manual for RevMan (Cochrane Collaboration, Copenhagen, Denmark) specifically noted that medians are not suitable for pooling because they preclude the derivation of means and SDs. Therefore, the 2 trials using this form of reporting results could not be included in this meta-analysis.

RevMan analyses of the trials showed favorable summary ES for the prevention trials ($n = 6$) and marginal ES for treatment studies ($n = 4$). The former evaluated probiotics' effect on PAD incidence after intervention, specifically at 2 years, in 1581 participants (786 probiotics, 795 placebo). The summary ES by the FE and RE models were quantified as relative RR of 0.69 (0.57, 0.83) and 0.66 (0.49, 0.89), respectively, favoring the use of probiotics as primary intervention for PAD. Interestingly, with the exclusion of the 1 prevention trial which, unlike the other 5,

administered probiotics only postnatally, this relative RR decreased to 0.61 by both FE and RE models, with minor differences in the upper limit of the 95% CI (0.76 vs 0.75, respectively).

The 4 treatment trials evaluated 299 participants (153 probiotics, 146 placebo). SCORAD scores were analyzed in terms of intergroup (final SCORAD scores) and intragroup (difference between preintervention and postintervention SCORAD scores) change. The ES for intergroup difference in postintervention SCORAD score was calculated as -6.64 points ($-9.78, -3.49$) and -8.56 ($-18.39, 1.28$), respectively, in the probiotics group. The FE and RE model-based ES were even less apparent for intragroup SCORAD change, with a greater reduction of SCORAD score throughout the intervention by -1.06 ($-3.86, 1.73$) and -1.37 ($-4.81, 2.07$) in the probiotics group.

DISCUSSION

Meta-analyses of 6 prevention and 4 treatment clinical trials indicate that current evidence favors the use of probiotics for prevention but not for treatment. There was a significant risk reduction by as much as 61% associated with the use of prenatal and/or postnatal probiotics for PAD prevention. An additional analysis, which excluded the 1 study with a purely postnatal protocol, revealed a lower relative RR with both FE and RE methods. This suggests that a prenatal component may be clinically important for maximizing probiotics' prophylactic potential.

In terms of treatment, the summary ES derived for both intergroup and intragroup differences failed to reach statistical significance. By the FE method, the final SCORAD difference was calculated as a weighted mean difference (WMD) of -6.64 ($-9.78, -3.49$), but this significance disappeared with the RE method. Regardless of which model was used, the WMD for intragroup difference, that is, preintervention and postintervention SCORAD change, failed to reach statistical significance.

Despite the disappointing conclusion reached for treatment, one must take into account that the treatment studies tended to be smaller and more heterogeneous in study design than the prevention trials. Therefore, the qualitative assessment has greater bearing on the lessons learned from meta-analyzing the treatment trials. The observations gleaned from the scoring component of this study have generated the following recommendations for higher-quality treatment trials of probiotics that can potentially yield results appropriate for meta-analysis.

1. Inclusion criteria of stable AD

Two of the earliest treatment trials by Majamaa et al¹¹ and Isolauri et al¹² observed significant improvement in SCORAD scores in the probiotics groups, but also observed delayed improvement in the placebo group several weeks after the trial. None of the 9 trials detected significant intergroup difference in posttrial SCORAD scores. Two trials did, however, report a significantly greater percentage of participants with AD improvement in the probiotics versus the placebo group.^{14,15} Neither of these studies defined a threshold change in score *a priori* to be considered a meaningful clinical response.

Sistek et al¹⁶ defined normal variation of stable AD as less than or equal to 11 points on the SCORAD scale and monitored scores during the trial run-in period, because AD stability was part of the inclusion criteria. The criterion of Sistek et al¹⁶ for stable disease (ie, SCORAD change of 11 points or less during a study's run-in period) and the criterion of Brouwer et al¹⁷ for meaningful

Cho & Bero MQ Rankings

Primary Prevention Studies

Study	Reviewer 1	Reviewer 2	Quartiles
20. Kukkonen, Savilahti, et al (2007)	1.000	0.957	1st vs 1st
21. Abrahamsson (2007)	1.000	0.915	1st vs 1st
19. Taylor, Dunstan, et al (2007)	0.957	0.936	1st vs 1st
4. Kalliomäki, Salminen, et al (2001)	0.933	0.844	1st vs 2nd
9. Kalliomäki, Salminen, et al (2003)	0.892	0.846	2nd vs 2nd
7. Rautava, Kalliomäki, et al (2002)	0.822	0.809	2nd vs 2nd

Blue	Clinical
Red	Laboratory
Green	Clinical & Laboratory
<i>Italics</i>	Retrospective Analysis
**	Observational

Secondary Prevention Studies

Study	Reviewer 1	Reviewer 2	Quartiles
18. Sistik, Kelly, et al (2006)	0.979	0.894	1st vs 1st
14. Viljanen, Savilahti, et al (2005)	0.936	0.915	1st vs 1st
17. Brouwer, Wolt-Plompen, et al (2006)	0.915	0.809	2nd vs 2nd
16. Weston, Halbert, et al (2005)	0.878	0.872	2nd vs 1st
15. Viljanen, Kuitunen, et al (2005)	0.830	0.809	2nd vs 2nd
13. Viljanen, Pohjavuori, et al (2005)	0.809	0.787	3rd vs 3rd
12. Rosenfeldt, Benfeldt, et al (2004)	0.809	0.766	3rd vs 3rd
3. Isolauri, Arvola, et al (2000)	0.800	0.756	3rd vs 3rd
5. Kankaanpää, Yang, et al (2002)	0.791	0.778	3rd vs 3rd
6. Kirjavainen, Arvola, et al (2002)	0.767	0.733	4th vs 4th
10. Rosenfeldt, Benfeldt, et al (2003)	0.702	0.766	4th vs 3rd
11. Kirjavainen, Salminen, et al (2003)	0.667	0.702	4th vs 4th

Miscellaneous Clinical Studies

Study	Reviewer 1	Reviewer 2	Quartiles
8. Lodinová-Zádníková, Cukrowska, et al (2003)	0.724	0.634	4th vs 4th
2. Pessi, Sütas, et al (2000) **	0.410	0.615	4th vs 4th

FIG 3. Ranking list of prevention and treatment clinical trials using Cho & Bero's MQ scale. Fisher analysis of ranking results from the 4 quality scales indicated that this scale had the best interobserver and interscale reliability. The quartile status allows gauging of the relative quality of these 2 interventional approaches. Boldface studies are those that were meta-analyzed.

therapeutic response (ie, intragroup SCORAD change of 15 points) provide important points of reference for interpreting study results.

2. Inclusion criteria of IgE-dependent AD for probiotic treatment trials and/or *a priori* subgroup stratification on the basis of IgE status

The sensitization status of the host has been observed to influence the response to probiotics. One study demonstrated that LGG administration, along with milk, was associated with immunostimulatory effects in healthy hosts, whereas in milk-hypersensitive infants, it was associated with anti-inflammatory effects.¹⁸ More importantly, allergic sensitization/IgE-associated status and greater PAD severity correlated with stronger therapeutic response to probiotics supplementation. Four trials detected significant therapeutic benefit in sensitized participants with subgroup analysis.^{13,15,16,19} Similarly, 2 trials detected a greater response in participants with more severe PAD.^{14,16} Weston et al¹⁴ observed that slightly older children (mean age, 11.5 mo) with a more severe dermatitis (mean SCORAD = 41) were more likely (92%) to show improvement of lesions statistically with probiotics. These findings highlight the importance of

distinguishing between IgE-dependent and independent PAD and defining threshold PAD severity for future studies.

A similar assertion has been made for intrinsic factors that may be predictive of the preventative value of probiotics. Rautava et al,²⁰ on the basis of their study findings, argued that infants most likely to benefit from probiotics are those with elevated cord blood IgE concentration. However, this has limited applications because this particular study involved exclusively breast-feeding mothers who ingested the probiotics and presumably passed it to their infants through breast milk. One cannot discount the oft-cited window of opportunity of the prenatal phase.

3. Verification of probiotic quality and minimization of confounders

Clinical trials should verify the quality of the probiotics they use and explicitly state that the products fulfill the criteria that have been set by European Union-funded research groups: (1) of human origin, (2) demonstrate nonpathogenic behavior, even in immunocompromised hosts; (3) exhibit resistance to technological processes; (4) prove resistant to gastric acid and bile; (5) adhere to gut epithelial tissue and be able to persist, albeit for short

periods, in the gut; (6) able to produce antimicrobial substances and modulate immune responses; and (7) able to influence metabolic activities (eg, cholesterol assimilation, lactase activity, vitamin production).²¹

Verification of the quality of the probiotic may help minimize the presumed confounding effects of possible modulators of probiotics' allergy-alleviating effects. For instance, topical corticosteroids, antigen elimination diets, and concomitant antibiotics use have been cited as potential confounders in assessing probiotics' secondary therapeutic effects. Antigen elimination diets have shown clinical benefit when combined with probiotics. One study observed a restorative effect of an antigen elimination diet on intestinal barrier function and anti-inflammatory action (ie, decrease in fecal α -1-antitrypsin and TNF- α concentrations in infants receiving LGG) as well as significant clinical improvement in the extent, intensity, and subjective score for PAD.¹¹ A 2-week lead-in period has been suggested as a way of minimizing the confounding effects of improved adherence to previous therapeutic regimens, such as topic steroids and antigen elimination diets, associated with participation in a clinical trial.

Antibiotics use has been thought to antagonize the effects associated with probiotics use, with the rationale based on the fact that antibiotics are among several factors known to interfere with postnatal colonization of gut microflora, of which many probiotic products are members.²² On the whole, the confounding potential of concomitant antibiotics use appears to be low, but it cannot be completely discounted. One study conducted a *post hoc* secondary analysis after excluding infants who received antibiotics between first and third visits. This enhanced the detected PAD improvement in the probiotics group. Subsequent calculation showed SCORAD decreases to be of greater magnitude in the 2 probiotics groups than placebo. This secondary analysis, however, had no effect on calculated SCORAD trends for placebo.¹³ Furthermore, one must take into account that this was a *post hoc* secondary analysis and therefore cannot be given too much weight because interpretation is prone to error (ie, *post hoc* secondary analyses are observational by nature and not based on randomized comparisons).

Prenatal and postnatal administration of probiotics appears to be a safe intervention for pregnant women with a personal or familial atopic history that may confer protection against AD for their infants. Pooled data from 6 published trials indicate that it significantly reduced the likelihood of having PAD at age 2 years. Evidence for the therapeutic value of probiotics for established AD is equivocal. Although statistically significant SCORAD score change and intergroup difference have been reported by individual trials, summary effect sizes fell far from the cutoff of 15 SCORAD points set by Brouwer et al¹⁷ as corresponding to meaningful clinical change in severity as opposed to a normal variable course of PAD. Further studies of probiotics for PAD treatment are needed that reflect the recommendations generated from this qualitative and quantitative review of the current literature.

We are indebted to Drs Julian Crane, Susan Prescott, and Vibeke Rosenfeldt for providing primary data to enable meta-analysis and Dr Michael Brimacombe for guiding and reviewing the statistical analyses of the quality scale assessments.

Clinical implications: A prenatal and postnatal regimen of probiotics is an effective option for pregnant women with risk factors for atopic disease to help prevent persistent PAD in their children.

REFERENCES

1. Isolauri E. Dietary modification of atopic disease: use of probiotics in the prevention of atopic dermatitis. *Curr Allergy Asthma Rep* 2004;4:270-5.
2. Prescott SL, Björkstén B. Probiotics for the prevention or treatment of allergic diseases. *J Allergy Clin Immunol* 2007;120:255-62.
3. Boyle RJ, Tang ML. The role of probiotics in the management of allergic disease. *Clin Exp Allergy* 2006;36:568-76.
4. Rautava S, Kalliomäki M, Isolauri E. New therapeutic strategy for combating the increasing burden of allergic disease: Probiotics-A Nutrition, Allergy, Mucosal Immunology and Intestinal Microbiota (NAMI) Research Group report. *J Allergy Clin Immunol* 2005;116:31-7.
5. Ogden NS, Bielory L. Probiotics: a complementary approach in the treatment and prevention of pediatric atopic disease. *Curr Opin Allergy Clin Immunol* 2005;5:179-84.
6. Yusuf S. Meta-analysis of randomized trials: looking back and looking ahead. *Control Clin Trials* 1997;18:594-601; discussion 61-6.
7. Helfenstern U. Data and models determine treatment proposals: an illustration from meta-analysis. *Postgrad Med J* 2002;78:131-4.
8. Cho M, Bero L. Instruments for assessing the quality of drug studies published in the medical literature. *JAMA* 1994;272:101-4.
9. Jadal AR, Moore AR, Carroll D, Jenkinson C, Reynolds DJ, Gavanagh DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1-12.
10. Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. *J Clin Epidemiol* 1998;51:1235-41.
11. Majamaa H, Isolauri E. Probiotics: a novel approach in the management of food allergy. *J Allergy Clin Immunol* 1997;99:179-85.
12. Isolauri E, Arvola T, Sutas Y, Moilanen E, Salminen S. Probiotics in the management of atopic eczema. *Clin Exp Allergy* 2000;30:1604-10.
13. Viljanen M, Savilahti E, Haahtela T, Juntunen-Backman K, Korpela R, Poussa T, et al. Probiotics in the treatment of atopic eczema/dermatitis syndrome in infants: a double-blind placebo-controlled trial. *Allergy* 2005;60:494-500.
14. Weston S, Halbert A, Richmond P, Prescott SL. Effects of probiotics on atopic dermatitis: a randomized controlled trial. *Arch Dis Child* 2005;90:892-7.
15. Rosenfeldt V, Benfeldt E, Nielsen SD, Michaelsen KF, Jeppesen DL, Valerius NH, et al. Effect of probiotic Lactobacillus strains in children with atopic dermatitis. *J Allergy Clin Immunol* 2003;111:389-95.
16. Sistiak D, Kelly R, Wickens K, Stanley T, Fitzharris P, Crane J. Is the effect of probiotics on atopic dermatitis confined to food sensitized children? *Clin Exp Allergy* 2006;36:629-33.
17. Brouwer ML, Wolt-Plompen SA, Dubois AE, van der Heide S, Jansen DF, Hoijer MA, et al. No effects of probiotics on atopic dermatitis in infancy: a randomized placebo-controlled trial. *Clin Exp Allergy* 2006;36:899-906.
18. Peltto L, Isolauri E, Lilius EM, Nuutila J, Salminen S. Probiotic bacteria down-regulate the milk-induced inflammatory response in milk-hypersensitive subjects but have an immunostimulatory effect in healthy subjects. *Clin Exp Allergy* 1998;28:1474-9.
19. Kukkonen K, Savilahti E, Haahtela T, Juntunen-Backman K, Korpela R, Poussa T, et al. Probiotics and prebiotic galacto-oligosaccharides in the prevention of allergic diseases: a randomized, double-blind, placebo-controlled trial. *J Allergy Clin Immunol* 2007;119:192-8.
20. Rautava S, Kalliomäki M, Isolauri E. Probiotics during pregnancy and breast-feeding might confer immunomodulatory protection against atopic disease in the infant. *J Allergy Clin Immunol* 2002;109:119-21.
21. O'Sullivan GC, Kelly P, O'Halloran S, Collins C, Collins JK, Dunne C, Shanahan F. Probiotics: an emerging therapy. *Curr Pharm Des* 2005;11:3-10.
22. Cucchiara S, Falconieri P, Di Nardo G, Parcelii MA, Dito L, Grandinetti A. New therapeutic approach in the management of intestinal disease: probiotics in intestinal disease in paediatric age. *Dig Liv Dis* 2002;34:S44-7.

Kappa = Reliability/Agreement

Inter-Observer Agreement

Cho & Bero MQ

n = 21		Reviewer 1				
Reviewer 2	Quartile	1st	2nd	3rd	4th	
	1st	5	0	0	0	0
	2nd	0	5	0	0	0
	3rd	0	0	4	1	1
	4th	0	0	1	5	5

Unweighted Kappa = 0.8727 (0.7041 - 1, 95% CI)
 Linear Weighted Kappa = **0.9250** (0.8245 - 1, 95% CI)
 Quadratic Weighted Kappa = 0.9632 (no CI calculated)

Jadad

n = 19		Reviewer 1				
Reviewer 2	Score	1.00	0.75-0.99	0.50-0.74	0.25-0.49	
	1.00	3	0	0	0	0
	0.75-0.99	3	2	2	0	0
	0.50-0.74	0	4	1	1	1
	0.25-0.49	0	0	1	2	2

Unweighted Kappa = 0.2201 (0 - 0.5475, 95% CI)
 Linear Weighted Kappa = **0.4890** (0.2350 - 0.7430, 95% CI)
 Quadratic Weighted Kappa = 0.7195 (no CI calculated)

Cho & Bero CR

n = 21		Reviewer 1				
Reviewer 2	Score	1.00	0.90-0.99	0.80-0.89	0.70-0.79	
	1.00	2	3	3	2	2
	0.90-0.99	0	2	1	1	1
	0.80-0.89	1	2	1	2	2
	0.70-0.79	0	0	1	0	0

Unweighted Kappa = 0.0175 (0 - 0.3274, 95% CI)
 Linear Weighted Kappa = **0.0506** (0 - 0.2815, 95% CI)
 Quadratic Weighted Kappa = 0.0870 (no CI calculated)

Delphi "Don't Know"

n = 19		Reviewer 1				
Reviewer 2	Score	Zero	One	Two	Three	
	Zero	1	2	2	1	1
	One	1	2	0	0	0
	Two	0	2	5	1	1
	Three	0	0	1	1	1

Unweighted Kappa = 0.2720 (0 - 0.5813, 95% CI)
 Linear Weighted Kappa = **0.3179** (0.0219 - 0.6139, 95% CI)
 Quadratic Weighted Kappa = 0.3410 (0 - 0.7236)

FIG E1. Kappa values indicate that the Cho & Bero MQ scale is superior to the other 3 scales in terms of interobserver and interscale concordance.

Inter-Scale Agreement

Cho & Bero MQ vs. Cho & Bero CR

REVIEWER 1

n = 21

CR	MQ			
	Quartile	1st	2nd	3rd
	1.00	0	2	0
	0.90-0.99	3	1	2
	0.80-0.89	2	2	0
	0.70-0.79	0	0	3

Unweighted Kappa < 0

Linear Weighted Kappa < 0

Quadratic Weighted Kappa < 0

Average Linear Weighted Kappa Value = not calculated (< 0.0488)

REVIEWER 2

n = 21

CR	MQ			
	Quartile	1st	2nd	3rd
	1.00	3	2	2
	0.90-0.99	1	0	1
	0.80-0.89	1	3	2
	0.70-0.79	0	0	0

Unweighted Kappa = 0.0597 (0 - 0.3550, 95% CI)

Linear Weighted Kappa = **0.0488** (0 - 0.3272, 95% CI)

Quadratic Weighted Kappa = 0.0406 (no CI calculated)

Cho & Bero MQ vs. Jadad

REVIEWER 1

n = 19

Jadad	MQ			
	Quartile	1st	2nd	3rd
	1.00	4	2	0
	0.75-0.99	1	2	2
	0.50-0.74	0	1	1
	0.25-0.49	0	0	2

Unweighted Kappa = 0.223 (0 - 0.5150, 95% CI)

Linear Weighted Kappa = **0.4771** (0.2475 - 0.7067, 95% CI)

Quadratic Weighted Kappa = 0.6863 (0.5702 - 0.8024, 95% CI)

Average Linear Weighted Kappa Value = 0.5305

REVIEWER 2

n = 19

Jadad	MQ			
	Quartile	1st	2nd	3rd
	1.00	2	1	0
	0.75-0.99	3	3	1
	0.50-0.74	0	1	3
	0.25-0.49	0	0	1

Unweighted Kappa = 0.3643 (0.0581 - 0.6705, 95% CI)

Linear Weighted Kappa = **0.5839** (0.3724 - 0.7954, 95% CI)

Quadratic Weighted Kappa = 0.7717 (no CI calculated)

FIG E1. (Continued).

Cho & Bero MQ vs. Delphi "Don't Know"

REVIEWER 1

n = 19

MQ

Delphi

Quartile	1st	2nd	3rd	4th
Zero	1	1	0	0
One	3	2	0	1
Two	1	1	4	2
Three	0	1	1	1

Unweighted Kappa = 0.2230 (0 - 0.5202, 95% CI)

Linear Weighted Kappa = **0.3416** (0.0767 - 0.6065, 95% CI)

Quadratic Weighted Kappa = 0.4722 (0.0036 - 0.9408, 95% CI)

Average Linear Weighted Kappa Value = 0.4933

REVIEWER 2

n = 19

MQ

Delphi

Quartile	1st	2nd	3rd	4th
Zero	4	1	1	0
One	1	2	0	0
Two	0	2	4	2
Three	0	0	0	2

Unweighted Kappa = 0.5037 (0.2123 - 0.7951, 95% CI)

Linear Weighted Kappa = **0.6449** (0.4073 - 0.8825, 95% CI)

Quadratic Weighted Kappa = 0.7672 (0.4903 - 1.0, 95% CI)

Jadad vs. Cho & Bero CR

REVIEWER 1

n = 19

Jadad

CR

Quartile	1.00	0.75-0.99	0.50-0.74	0.25-0.49
1.00	1	1	1	0
0.90-0.99	2	3	1	0
0.80-0.89	3	1	0	1
0.70-0.79	0	1	2	2

Unweighted Kappa = 0.0919 (0 - 0.3786, 95% CI)

Linear Weighted Kappa = **0.2227** (0 - 0.5125, 95% CI)

Quadratic Weighted Kappa = 0.3770 (0 - 0.8025, 95% CI)

Average Linear Weighted Kappa Value = not calculated (< 0.2227)

REVIEWER 2

n = 19

Jadad

CR

Quartile	1.00	0.75-0.99	0.50-0.74	0.25-0.49
1.00	2	3	2	2
0.90-0.99	0	1	2	0
0.80-0.89	1	3	1	1
0.70-0.79	0	0	1	0

Unweighted Kappa < 0

Linear Weighted Kappa < 0

Quadratic Weighted Kappa < 0

FIG E1. (Continued).

Jadad vs. Delphi "Don't Know"

REVIEWER 1

n = 19

= 19		Jadad				
		Quartile	1.00	0.75-0.99	0.50-0.74	0.25-0.49
Delphi	Zero	2	0	0	0	
	One	3	1	2	0	
	Two	0	4	2	2	
	Three	1	1	0	1	

Unweighted Kappa = 0.0919 (0 - 0.4169, 95% CI)
Linear Weighted Kappa = **0.2621** (0 - 0.5509, 95% CI)
Quadratic Weighted Kappa = 0.3838 (0 - 0.8302, 95% CI)

Average Linear Weighted Kappa Value = 0.3216

REVIEWER 2

n = 19

n = 19		Jadad				
		Quartile	1.00	0.75-0.99	0.50-0.74	0.25-0.49
Delphi	Zero		2	3	1	0
	One		1	2	0	0
	Two		0	2	3	3
	Three		0	0	2	0

Unweighted Kappa = 0.1493 (0 - 0.4573, 95% CI)
Linear Weighted Kappa = **0.3810** (0.1632 - 0.5988, 95% CI)
Quadratic Weighted Kappa = 0.5980 (0.2987 - 0.8973, 95% CI)

Delphi "Don't Know" vs. Cho & Bero CR

REVIEWER 1

n = 19

19	Delphi				
	Quartile	Zero	One	Two	Three
CR	1.00	1	1	1	0
	0.90-0.99	0	3	2	1
	0.80-0.89	1	1	1	2
	0.70-0.79	0	1	4	0

Unweighted Kappa < 0
Linear Weighted Kappa **< 0**
Quadratic Weighted Kappa < 0

Average Linear Weighted Kappa Value = not calculated (< 0.1259)

REVIEWER 2

n = 19

= 19		Delphi				
		Quartile	Zero	One	Two	Three
CR	1.00	3	2	4	0	
	0.90-0.99	1	0	1	1	
	0.80-0.89	2	1	3	0	
	0.70-0.79	0	0	0	1	

Unweighted Kappa = 0.0806 (0 - 0.4240, 95% CI)
Linear Weighted Kappa = **0.1259** (0 - 0.4983, 95% CI)
Quadratic Weighted Kappa = 0.2073 (0 - 0.5829, 95% CI)

FIG E1. (Continued).

Inter-Observer & Average Inter-Scale Agreement

(Summary with Linear Weighted Kappa Values)

	MQ	CR	Jadad	Delphi
MQ	0.9250			
CR	(< 0.0488)	0.0506		
Jadad	0.5305	(< 0.2227)	0.4890	
Delphi	0.4933	(< 0.1259)	0.3216	0.3179

	Qualitative Value	
Kappa Value	Sackett (1991)	Altman (1991)
< 0	Poor	
0 - 0.20	Slight	Poor
0.21 - 0.40	Fair	Fair
0.41 - 0.60	Moderate	Moderate
0.61 - 0.80	Substantial	Good
0.81 - 1.0	Almost Perfect	Very Good

So objectively speaking, what is the definition of a raw score of 0.90 or 0.80? How much "better" is the former than the latter?

Jadad has been something of an industry standard among meta-analysis study inclusion measures.

However, according to our statistical analyses, it doesn't quite merit the title of "gold standard."

On the contrary, the Cho & Bero MQ appears to be superior both in terms of inter-observer and inter-scale agreement.

FIG E1. (Continued).

TABLE E1. Quality scales for trial scoring

Scale	Questions
Cho & Bero's MQ Scale	Point system: yes = 2; partial = 1; no or not applicable = 0 1. Study design (1 of 15 choices); case report = 1 point; time series or uncontrolled = 2 points; cohort or case control = 3 points; unrandomized controlled trial = 4 points; randomized control trial = 5 points 2. What was the study question? (short answer) 3. Was the study question sufficiently described? 4. Was the study design appropriate to answer the study question? 5. Were both inclusion and exclusion criteria specified? (If case study, check not applicable) 6. For case studies only: Were patient characteristics adequately reported? 7. Were subjects appropriate to the study question? 8. Were the control subjects appropriate? (If no controls were used, check no) 9. Were subjects randomly selected from the target population? 10. If subjects were randomly selected, was the method of random selection sufficiently well described? 11. If subjects were randomly allocated to treatment groups, was the method of randomization described? (If subjects not randomly allocated, check not applicable) 12. If blinding of investigators to intervention was possible, was it reported? 13. If blinding of subjects to intervention was possible, was it reported? 14. Was measurement bias accounted for by methods other than blinding? 15. Were known confounders accounted for by study design? 16. Were known confounders accounted for by analysis? 17. Was there a sample size justification before the study? 18. Were <i>post hoc</i> power calculations or CIs reported for statistically nonsignificant results? 19. Were statistical analyses appropriate? 20. Were the statistical tests stated? 21. Were exact <i>P</i> values or CIs reported for each test? 22. Were attrition of subjects and reason for attrition recorded? 23. For those subjects who completed the study, were results completely reported? 24. Do the findings support the conclusions?
Cho & Bero's CR Scale	Point system: yes = 2; partial or insufficient evidence = 1; no or no controls = 0 1. Were the therapeutic outcomes measured in the study important? 2. Were the subjects of the study representative of patients who would actually use the drug? 3. Was the comparison group clinically meaningful? 4. Was the treatment effect clinically meaningful? 5. Were the side effects adequately measured? 6. Was approval from an institutional review board explicitly reported? 7. As far as could be determined from the article, was the study ethical?
Jadad Scale	Point system: yes = 1; no = 0 Additional 1 point: • Given if method of randomization and/or double blinding is described and it was appropriate • Deducted if method of randomization is described but is inappropriate or if study claimed to be double blind but method was inappropriate (eg, comparison of tablet vs injection) 1. Was the study described as randomized (this includes the use of words such as randomly, random, and randomization)? 2. Was the study described as double-blind? 3. Was there a description of withdrawals and dropouts?
Delphi Scale	Answer choices: yes, no, don't know Each question has 3 answer choices: yes, no, and don't know, which allows for use by nonexperts of epidemiology and trained readers of scientific literature. There is no point system given. 1. Treatment allocation; (a) Was a method of randomization performed? (b) Was the treatment allocation concealed? 2. Were the groups similar at baseline regarding the most important prognostic indicators? 3. Were the eligibility criteria specified? 4. Was the outcome assessor blind? 5. Was the care provider blind? 6. Was the patient blind? 7. Were point estimates and measures of variability presented for the primary outcome measures? 8. Did the analysis include an intention-to-treat analysis?

TABLE E2. Summary of clinical studies of probiotics on PAD: primary prevention studies (double-blind randomized controlled trials)

Study	No. of infants	Strain	Frequency and duration	Results
Kalliomaki et al (2001)	134	LGG	1×10^{10} cfu/g $\times$ 2 capsules (to mothers 2-4 wk prenatally; to infants 24 wk postnatally)	At 2 years, AD frequency halved in infants given probiotics: 23% (16 of 64) vs 46% (31 of 68) of placebo ($P = .008$); RR = 0.51 (2.6-15.6); number needed to treat = 4.5
Rautava et al (2002)	57	LGG	2×10^{10} cfu/g daily (prenatal and 3 mo postnatally, to breast-feeding mothers)	Maternal intake of probiotics ($n = 27$) associated with reduction in AD prevalence compared with placebo ($n = 30$)
Kalliomaki et al (2003)	107	LGG		Four-year follow-up of study by Kalliomaki et al (2001); protective effect against AD still detected: RR = 0.57 in LGG, compared with placebo
Kukkonen et al (2007)	925	Four strains of probiotics with a prebiotic	Mothers for 2-4 wk before delivery and 6-mo supplementation of infants	No effect of probiotics ($n = 461$) vs placebo ($n = 464$) on cumulative incidence of allergic diseases, but tended to reduce IgE-associated (atopic) diseases: OR = 0.71 (0.50, 1.00); $P = .052$ Probiotics reduced eczema (OR = 0.74 [0.55, 0.98]; $P = .035$) and atopic eczema (OR = 0.66 [0.46, 0.95]; $P = .025$)
Taylor et al (2007)	178	<i>Lactobacillus acidophilus</i> LAVRI-A1	Newborns of women with allergy ($n = 231$) received either probiotics (3×10^9 cfu) or placebo daily for 6 mo	No reduction in AD at 6 months (probiotics [$n = 23$ of 89; 25.8%] versus placebo [$n = 20$ of 88; 22.7%] groups [$P = .629$]) or at 1 year (probiotics [38/88; 43%] vs placebo [34/87; 39%] groups [$P = .581$]) Paradoxical increase in sensitization to allergens in probiotics-supplemented infants; calculated OR = 1.184; calculated 95% CI = 0.6483, 2.1648
Abrahamsson et al (2007)	188	<i>Lactobacillus reuteri</i> ATCC 55730 (10 cfu)	10^8 cfu Daily to mothers from gestational age 36 wk to delivery, then infants supplemented for 1 y, then followed for 1 y	Cumulative incidence of eczema similar (36% [$n = 34$] in probiotics group vs 34% [$n = 32$] in placebo group) Probiotics group had less IgE-associated eczema during second year (8% vs 20% [$P = .02$]); adjusted OR = 0.36; $P = .047$

cfu, Colony-forming units; OR, odds ratio.

TABLE E3. Summary of clinical studies of probiotics on PAD: secondary prevention studies (double-blind RCTs)

Study	Total no., placebo no., and probiotics no.	Probiotics strains used	Frequency and duration	Results
				Endpoints = intragroup and intergroup change in SCORAD (calculated using formula: $[A/5] + [3.5B] + C$; maximum score = 103); serum markers; both
Majamaa and Isolaure (1997)	Part I (RCT): 31 infants fulfilling Hanifin criteria 2.5-15.7 mo Part II (no control): 10 atopic infants 0.6- 8.5 mo breast-fed by mothers given probiotics	LGG	Part I: Given to infants 5×10^8 cfu/g Part II: Given to nursing mothers 2×10^{10} cfu/g twice daily 4 wk	After 4 wk of intervention, significant SCORAD score reduction in probiotic group (probiotic group) BUT comparable SCORAD was achieved in placebo group in 2 mo AD severity reduction in group WhGG was a result of the following: (a) reduction of the extent ($P = .004$); (b) intensity ($P = .05$); (c) subjective score (.01) Significant reduction in concentration of fecal α -1-antitrypsin ($P = .03$).
Isolaure et al (2000)	27 Infants who manifested AD during exclusive breast-feeding	LGG (3×10^8 cfu/g) <i>Bifidobacterium</i> <i>lactis</i> Bb-12 (1×10^9 cfu/g)	4 wk	At 8 wk, significant SCORAD change ($P =$.01) from initial median score of 16 (7-25) was observed in the following: • All (9/9) of the LGG group (median SCORAD: 1) • All (9/9) of the <i>B. lactis</i> group (median SCORAD: 0) • 44% (4/9) of the placebo group (median SCORAD: 13.4) However, at 24 wk, the median SCORAD was 0 in all groups
Kankaanpaa et al (2002)	Study pool of 15 exclusively breast- fed infants fulfilling Hanifin criteria	LGG (3×10^8 cfu/g) B Bb12 (1×10^9 cfu/g)	Varying durations	Significant changes in plasma lipid PUFA composition were detected: • In B Bb12 probiotic group, retention of elevated proportion of eicosapentaenoic acid and decreased proportion of MUFA observed • In LGG group, relative percentage of total MUFA increased and arachidonic acid in plasma neutral lipids decreased • Both groups influenced the proportions of n-3 PUFA in neutral lipids; both reduced the proportion of α-linolenic acid
Kirjavainen et al (2002)	13 of 21 Breast-fed infants* with early- onset AD	B Bb12 (1×10^9 cfu/g) in EHF Mean daily intake: 8×10^{10} cfu/kg body weight		Bifidobacterial supplementation prevented increase in number of bacteroides and <i>Escherichia coli</i> during weaning, specifically bacteroides ($P = .07$) and <i>E. coli</i> ($P = .02$) SCORAD drop in 50% of placebo group and 100% of probiotics group (Fisher exact, $P = .07$)
Rosenfeldt et al (2003)	Cross-over study 43 Unselected patients (42% male; mean age, 5.2 y; age range, 1-13 y)	<i>Lactobacillus</i> <i>rhamnosus</i> 19070-2 <i>L. reuteri</i> DSM 12246 (10^{10} cfu) given 2 $\times$ daily	18 wk (6 wk of treatment or placebo, washout period, then treatment or placebo)	No overall significant change in total SCORAD after treatment • Significant change of a cruder measure of improvement $\rightarrow$ 56% of patients who took probiotics experienced improvement of eczema, compared with 15% of patients who took placebo ($P = .001$) • Greater treatment response greater in patients with at least 1 positive skin prick test response and elevated IgE levels; the SCORAD score in this subgroup decreased ($P = .03$) • In probiotics group, sECP levels decreased ($P = .03$)

(Continued)

TABLE E3. (Continued)

Study	Total no., placebo no., and probiotics no.	Probiotics strains used	Frequency and duration	Results Endpoints = intragroup and intergroup change in SCORAD (calculated using formula: [A/5] + [3.5B] + C); maximum score = 103); serum markers; both
Kirjavainen et al (2003)	35 Infants Mean age, 5.5 mo Three groups: * Placebo, 8 * Viable LGG, 14 * Heat-inactivated LGG, 1	LGG (1×10^9 cfu)	Intervention duration varied from 0.4 to 45.3 wk; mean, 7.5 wk	Treatment with heat-inactivated LGG associated with adverse gastrointestinal symptoms and diarrhea, thus recruitment of patients stopped after pilot phase SCORAD scores decreased in all groups from: • 13 to 8 in placebo group • 19 to 5 in viable LGG group • 15 to 7 in heat-inactivated group LGG supplementation did not affect bacterial numbers, but because of the low power of the study and heterogeneous diets, the lack of changes in the microbiota must be interpreted with caution
Rosenfeldt et al (2004)	41 Children (mean age, 4.0 y; range, 1-13 y) with moderate and severe AD	<i>L. rhamnosus</i> <i>L. reuteri</i> DSM (10^{10} cfu)	18 wk	Crossover study design Endpoints: SCORAD, serum IgE, GI symptoms Small intestinal permeability (Lactulose- Mannitol test), measured from only 34% (14/41) of the study population
Viljanen et al (2005)	230 Infants with atopic eczema- dermatitis Sx and suspected CMA (from original pool of 252)	LGG MIX - 4 probiotic strains: LGG, <i>L. rhamnosus</i> , <i>Bifidobacterium breve</i> , <i>Propionibacterium</i> <i>freudenreichi</i> (5×10^9 cfu BID)	4 wk	Three subgroups 1. IgE-associated AD: significant changes in paired pretreatment and post-treatment inflammatory markers • $\uparrow$ CRP (but not significant in whole population) • $\uparrow$ IL-6 during LGG treatment (significant based on Mann-Whitney <i>U</i> test ($P =$.036), but no change seen in MIX, even though it contains the same amount of LGG • $\uparrow$ IL-10 in MIX group ($P = .016$) 2. IgE associated CMA: $\uparrow$ serum E-selectin levels after Tx 3. IgE independent AD, IL-10 increased in LGG ($P = .046$) and MIX ($P = .039$) groups more so than in placebo
Viljanen et al (2005)	Same as row above			No significant differences in SCORAD score reduction between probiotics groups and placebo; AEDS symptoms improved in all treatment groups, reflected in decreasing SCORAD from baseline to immediately after Tx and 4 weeks after Tx cessation Statistically significant beneficial effects of probiotics observed to be limited to IgE- associated AD, and this observed effect was enhanced after excluding the large proportion of infants treated with antibiotics during or shortly after the probiotics
Viljanen et al (2005)	Same as row above			Fecal inflammatory marker ECP levels decreased significantly in all groups Significant reduction of IgA in placebo group In IgE-associated CMA infants, adjusted postchallenge fecal IgA was greater with the LGG than the placebo

(Continued)

TABLE E3. (Continued)

Study	Total no., placebo no., and probiotics no.	Probiotics strains used	Frequency and duration	Results
				Endpoints = intragroup and intergroup change in SCORAD (calculated using formula: [A/5] + [3.5B] + C); maximum score = 103); serum markers; both
Weston et al (2005)	53 of 56 children age 6-18 mo with moderate or severe AD; all had modified SCORAD >25	<i>Lactobacillus</i> <i>fermentum</i> VRI-033 PCC, probiotics (1 × 10 ⁹ cfu/g) BID	8 weeks	Slightly older children (mean age, 11.5 mo) with more severe dermatitis (mean SCORAD = 41) are more likely (92%) to show improvement of lesions statistically with probiotics Significantly more children in probiotics group (n = 24; 92%) had score improvement than those given placebo (n = 8; 30%) at t = 16 wk
Brouwer et al (2006)	50 Infants under 5 mo of age Placebo, 17 <i>L rhamnosus</i> , 17 LGG, 16	<i>L rhamnosus</i> LGG (5 × 10 ⁹ cfu/100 mL formula)	12 wk	No significant clinical effect of probiotics on AD severity was detected Differences in SCORAD at randomization (placebo, 29.7; <i>L rhamnosus</i> , 24.23; LGG, 29.4) and subsequent decrease in SCORAD during treatment (placebo, 2.6 points/mo; <i>L.rh</i> , 2.65 points/mo; <i>L.GG</i> , 2.8 points/mo) were not significant
Sistek et al (2006)	59 Children Placebo, 30 Probiotics, 29	Combined <i>L rhamnosus</i> and <i>B lactis</i> (2 × 10 ¹⁰ cfu/g daily)	12 weeks	Significant AD improvement only observed in food-sensitized children (n = 19 of 29 in probiotics; n = 24 of 30 in placebo) SCORAD geometric mean ratio among food- sensitized children given probiotics was 0.73 (95% CI, 0.54-1.00; P = .047) Intention-to-treat analysis done at wk 12. Per- protocol analysis (dropouts not assessed) done at wk 16 (24 placebo and 25 probx analyzed)

PUFA, Polyunsaturated fatty acids; MUFA, monounsaturated fatty acids; EHF, extrahydrolyzed formula; sECP, serum eosinophilic cationic protein; CMA, cow's milk allergy; AEDS, atopic eczema-dermatitis syndrome.

*Hanifin and Rajka's criteria require 3 or more major features (eg, pruritus, personal or family history of atopy) and 3 or more minor features (eg, elevated serum IgE, xerosis, early age of onset).

TABLE E4. Summary of clinical studies of probiotics on PAD: miscellaneous clinical studies

Study	Study design	Sample size	Probiotics used, frequency, and duration	Results
Pessi et al (2000)	Observational study; no control group	9 Atopic children meeting the Hanifin criteria*	LGG (10^{10} cfu) BID 4 wk	Serum (IL-10) determined to differ significantly in before, early, and late samples ($P < .001$, ANOVA) detected in 2 of 9 patients at 4 wk and 7 of 9 patients at 8 wk Enhanced IL-10 production in PBMCs preceded rise in serum IL-10 Overexpression of IL-10 has been demonstrated spontaneously and after allergen challenge in atopic eczema
Lodinova-Zadnikova et al (2003)	Questionnaire-based retrospective study	77 Preterm infants and 150 full-term infants colonized 10 and 20 y ago, respectively	<i>E coli</i> serotype 083:K24:H31 A1: orally colonized 3 times during first week of life, each time with 1 mL of suspension containing 5×10^8 microbes. B1: orally colonized 3 times during first 4 weeks of life	20 y ago: groups A1 and A2 (full-term infants) 10 y ago: groups B1 and B2 (premature infants) • Survey return rates: A1 = 150 (31%); A2 = 144 (28%); B1 = 77 (40%); B2 = 55 (33%) After 10 y, there are significant differences in occurrence of allergies and repeated infections ($P < .01$) between individuals who were colonized and not colonized as premature infants After 20 y, there is a significant difference in allergies occurrence ($P < .01$) but not repeated infections in subjects born full-term infants