**Materials and Methods**

*Search Strategy*

The systematic review and meta-analysis were performed in accordance with MOOSE [21]. We conducted a comprehensive literature search of the following databases: PubMed (1966 to August 28, 2018), EMBASE (1974 to August 28, 2018), MEDLINE (1946 to August 28, 2018) and Cochrane Library (up to July 2018). The search terms or text keywords used were: (breast-feeding OR breast feeding OR breastfeeding OR breastfeed) AND (endogenous eczema OR infantile eczema OR atopic eczema OR atopic dermatitis). In addition, the reference lists of all included publications and relevant reviews were checked manually for potential eligible articles.

*Study Selection*

Two authors (R. Dai and L. Lu) identified eligible publications independently; disagreement was resolved with consensus among all authors. Publications evaluating the association of breastfeeding and AD were considered for inclusion if: (1) the study had a prospective cohort design; (2) the authors reported data from an original paper; (3) the authors reported risk estimates of AD by breastfeeding status; (4) duration of breastfeeding for at least 3 months; (5) maternal recall of the child’s feeding history up to the age of 12 months; (6) never breastfeeding (e.g., exclusive formula feeding or cow milk feeding) or partial feeding (breastfeeding less than 3 months) as the reference group; (7) specific AD assessment methods were provided by the authors; and (8) studies were published in English. Studies were excluded on the basis of the following criteria: (1) not original articles (e.g., review, meta-analysis, case report or guideline etc.); (2) not human studies (e.g. in vitro study); (3) studies reporting crude associations without any adjustment. When several studies were published based on the same data, only the most informative one (e.g., the most recent one or the study with the largest size) was included.

*Data Extraction*

For each included study, two authors (X. Fan and Y. Yu) independently extracted the following information using a prestructured data collection form: first author, year of publication, country location, subject baseline (sample size and male number), follow-up years, feeding behavior, feeding assessment, type of reference group, breastfeeding duration, outcome assessment (self-report, clinical examination or physician diagnosis), the criteria to diagnose AD, family history of atopy, adjusted variables, relative risks (RRs) with corresponding 95% confidence intervals (CIs). If some study did not provide overall RRs, but presented odds ratio (OR) for different AD status, we converted an OR to an RR according to Grant’s method [22]. If certain data were unavailable or unclear in published form, an e-mail contact to authors for original information was made.

*Quality Assessment*

The quality of each included study was assessed using the Newcastle-Ottawa Scale (NOS) [23], which addressed representativeness of the exposed population, selection of the nonexposure group, assessment of exposure and outcome, elimination of outcome at baseline, adequacy of follow-up, proportion of loss to follow-up and confounding adjustment (2 items). Overall, the scale scores ranged from 0 (lowest) to 9 (highest). Studies that met 5 or more of the NOS items were considered as high quality. We also compared our NOS scores with those reported by Lodge et al. [20] to make the assessment results more reliable.

*Statistical Analysis*

Since only prospective cohort studies were included in our meta-analysis, we used RRs with corresponding CIs to evaluate the association across studies. In order to control confounding variables, we adopted RRs adjusted for one or more confounders. Some studies did not provide overall RRs for breastfeeding but presented separate RRs for different duration levels of breastfeeding. For the reasons given above, we combined the corresponding risk estimates on the basis of Hamling’s method [24]. If a particular study reported several risk estimates of breastfeeding varieties using an independent reference category (e.g., independent estimates of children with and without atopic family history), we considered each variety as a separate study.

Separate meta-analyses were conducted to investigate the association between AD and breastfeeding (total and exclusive). The total breastfeeding comprised both breastfeeding and exclusive breastfeeding. The data of breastfeeding behavior were collected into two categories of breastfeeding and exclusive breastfeeding. If no other milk products, substitutes or solids were added into the infants’ diet at least for 3 months, the feeding behavior was included in the category of “exclusive breastfeeding.” If both the breast milk and other milk types were given to the infants, the study was included into the category of “breastfeeding.” If no detailed information regarding breastfeeding behavior was disclosed, the study was also included in the category of “breastfeeding.”

A random-effects model was used to estimate pooled RRs, which was more conservative than the fixed-effects one and allowed for between-study heterogeneity [25]. Heterogeneity across studies was evaluated by the χ2 statistic and quantified by *I2* [26, 27]. A study was considered as heterogeneous if *p* < 0.1 or *I*2> 50%. When substantial heterogeneity was present, subgroup analyses and metaregression were performed as exploratory analysis to evaluate the influences of selected factors and participant characteristics on the results. The predefined covariates included type of reference group (partial and never breastfeeding), breastfeeding duration, AD assessment methods, AD stages, family history of atopy, publication year, sample size, the length of follow-up and national income level. As for the reference group, breastfeeding or exclusive breastfeeding for less than 3 months or combined breastfeeding was defined as partial breastfeeding. If the reference group never drank any breast milk, we defined the group as never breastfeeding. The duration of breastfeeding was categorized into two groups as: 3–6 months and >6 months. Since different assessment methods of AD were reported in studies, we combined “physician diagnosis” and “clinical examination” into one subgroup for these methods were more precise. AD stages constituted three subgroups: the infant AD from birth to 2 years of age, the childhood AD between 2 and 12 years of age, and the adult AD of patients older than 12 years [28]. The level of national income was categorized on the basis of purchasing power parity (PPP). The country was regarded as high income if its gross domestic product (PPP) per capita (int$) was >35,000 according to the World Bank (2017) [29]. Also, sensitivity analyses were carried out to investigate the influence of an individual study on the pooled estimate by subsequently omitting studies and determining the potential weighted outliers.

Potential publication bias was evaluated by a funnel plot, Egger’s test [30] and Begg’s test [31]. The trim-and-fill method was also used to further assess the possible effect of publication bias [32]. All analyses were performed with Stata version 13.0 (Stata Corporation, College Station, TX, USA). The significance level was set to *p* < 0.05.