**Methods**

*Search*

We searched for epidemiological studies in the electronic databases PubMed, Web of Science, EMBASE, Cochrane Library, China National Knowledge Infrastructure (CNKI), Wanfang, and Chongqing VIP using the keywords “alopecia areata” in combination with the terms “thyroid disease”, “thyroiditis”, “thyroid antibody”, “hypothyroidism”, or “hyperthyroidism” in the title or abstract. All studies published from January 1980 to October 2017 were searched. The search languages were limited to English and Chinese.

*Inclusion Criteria and Data Extraction*

To satisfy the analysis requirements and to reduce selection deviation, the studies needed to meet the following criteria for inclusion: (1) all cases were confirmed by a dermatologist and not by a record or based on self-reports; (2) a case-control design was used; (3) sufficient data for cases and controls were provided to enable calculation of the odds ratio (OR) with a 95% confidence interval (CI) and *p* value. Reviews, case reports, and letters were excluded from this meta-analysis. The information was extracted from the included studies by two independent researchers. Any differences were resolved by discussion between the two researchers; if no consensus could be reached, another researcher was consulted. Another two independent reviewers, who were not blinded with regard to the author or journal, assessed the risk of bias among the studies included using the Newcastle-Ottawa Scale (NOS) [7].

*Statistical Analysis*

We employed a systematic analysis to calculate the risk of thyroid disorder in patients with AA based on the pooled OR with the corresponding 95% CI. Heterogeneity among studies was assessed using Cochran’s Q test and the *I*2 statistic. If the data showed low or moderate heterogeneity (*I*2 <50%), a fixed-effects model was used; otherwise, a random-effects model was used. Additionally, sensitivity analysis was performed to examine the influence of any particular study on the pooled estimate. Publication bias was evaluated using Egger’s test and Begg’s test. The significance level was set at a *p* value <0.05. All statistical analyses were performed using Stata version 12.0 (College Station, Texas, USA).