Reviewer #1
This is a well-written review article focusing on different diagnostic approaches and practice between Asia and western countries. The authors give a clear account of the current opinion and research status of clinicopathological characteristics, diagnostic modality especially ancillary techniques of primary thyroid lymphoma. Only minor revisions are suggested for the authors:

Comment 1: The following sentence in Abstract “Nearly 90% of all patients presented with early stage (stage I/II) disease, but a higher proportion of patients from the Western studies presented with higher stage (stage III/IV) disease (12.2% vs 3%)” seems unclear. Please clarify it.

Reply 1: The sentence has now been split into two statements and seems more clear now.

Comment 2: In page 2, the abbreviation “HT” should be added after the term “Hashimoto thyroiditis”.

Reply 2: This abbreviation has been added.

Reviewer #2
Comment 1: Table 2: It is unclear how the incidence of each parameter was calculated. The author should explain it in the text and describe them in more detail. For example, the prevalence of MALT lymphoma: 3.0% to 53.0% (median 36.3%) in Asian countries. The frequency is reported to total number is not meaningful.

Reply 1: Prevalence of specific PTL subtypes was calculated as proportion of number of cases of the subtype reported in literature to the total number of PTL cases reported. The prevalence of patients in a particular clinical stage was also calculated similarly. The same has been added to the manuscript.

Comment 2: On an ancillary study, the description is poor and not perfect. This reviewer recommends the author to cite the following research.

Reply 2: The recommended reference has been included in the study. The discussion has also been improved.
Comment 3: Table 3: The author shows the incidence of ancillary studies in Table 3. However, I think that the data are not reliable. The frequency of preoperative ultrasound examination is too low. According to the data, fine needle aspiration cytology is performed without ultrasound guidance in most institutions. I think that "no description" does not mean "no performance."

Reply 3: This is true that the use of preoperative ultrasound may not have been mentioned in most publications, hence the row pertaining to pre-operative ultrasound has been deleted from Table 3.

Reviewer #3
This study analyzed 2850 cases of primary thyroid lymphoma (PTL) reported in 40 publications (18 Asian and 22 Western). There were differences between Asian and Western cohorts in the histological type and stage of presentation of PTL. Mucosa-associated lymphoid tissue lymphoma being most prevalent (36.3%) among Asians while diffuse large B cell lymphoma (71.9%) in the Western population. Nearly 90% of all patients presented with early-stage (stage I/II) disease, but a higher proportion of patients from the Western studies presented with higher stage (stage III/IV) disease (12.2% vs. 3%). This meta-analysis is an excellent study to overview the current approach to this rare disease in the world.

Major comments:
Comment 1: The flow cytometry is a powerful tool for the diagnosis of PTL (references 36 and 38). At the same time, the specificity of ultrasound and fine-needle aspiration cytology was reported to be low (diagnostic rate was 42-59% with FNA cytology alone, and the positive predictive value was 63-65 % with US alone). There are two approaches to obtain samples for the flow cytometry, 1) cytology samples with FNA and 2) tissue samples by either core needle biopsy, wedge biopsy, or thyroidectomy. Which sampling method did the author find more popular in PTL practice? If possible, please add any merits and demerits of both cytology samples and tissue samples.

Reply1: Majority of the studies used cytology material for the Flowcytometry except for two studies where resected material was used for flowcytometry. This has now been discussed in the manuscript also.

Minor comments:
Comment 2: On page 6, the author stated that none of the studies had used TBSRTC for categorizing aspirates. This reviewer believes that it was because most of them were studies before 2009 when the TBSRTC was first introduced. Only an exceptional study by Kakkar et al., which was published in 2019, used the Bethesda reporting system on cases diagnosed after 2009.

Reply 2: This could be the possible reason for lack of TBSRTC data from these stud-
ies, as except for very few studies majority were conducted before 2009. This has been added to the manuscript.