Dear Dr Weber,

Thank you for the opportunity to respond to the reviewer's comments. We are very grateful to the reviewer for his helpful suggestions which, we believe, have improved our manuscript. The manuscript has been revised by explaining further the structure of the hierarchy used in the statistical model in order to group the countries; adding limitations in the Discussion regarding the potential influence of the countries group hierarchy to the income related patterns of the findings; and by highlighting further countries with estimates obtained from observed data and those extrapolated.

Below we have provided a point-by-point response to each of the reviewer's comments. All the changes have been tracked in the text and table of the revised manuscript.

Yours sincerely,

Rosa Parisi

Reviewer: 1

The authors now more clearly report the methods which they have followed in producing this report. It clearly has been a substantial project and the authors should be congratulated on their work.

However, it is now clear that the reported findings largely depend on the model which has been fitted to the data, particularly for the countries and regions where data are sparse or absent. The structure of the model (in terms of the grouping of countries into regions and regions in super regions) will have had a strong influence on the reported results which needs consideration and discussion. It is also important to be clear as to what results are supported by observed data, and which are extrapolations – currently all results are presented as being equally valid, which is not the case.

A: We are grateful to the reviewer for his useful comments, which we have now addressed. We have provided additional explanation about the structure of the hierarchy used in the statistical model in order to group the countries; added limitations in the Discussion regarding the potential influence of the countries group hierarchy to the income related patterns of the findings; and highlighted further which countries and regions' estimates were obtained from observed data and which were extrapolated.
Furthermore, in order to facilitate access and readability of the supplementary information, the Supplementary Material has now been split into two documents: Supplementary Material 1, which includes all the relevant information relating to the systematic review (stage I); and Supplementary Material 2, which includes all the information relating to the statistical model and meta-regression analysis (stage II).

1. **Whilst the use of the hierarchical model is appropriate, the interpretation of results needs to be made with care. As was raised in the previous comments, there are concerns that the level of extrapolation that is being used by the model where data are missing for particular countries, regions or super-regions is inadequately acknowledged by the authors, and is not suitably flagged to a reader. I would expect to read more about the impact of the degree of extrapolation in the limitations section of the discussion.**

A: Studies identified from the systematic review and included in the statistical model came from 19% of the countries of the world. As mentioned by the reviewer, the hierarchical model will help in dealing with the sparseness of the data. In addition, it is important to clarify that:

- *i) Given the structure of the hierarchy/geographical groups, there are countries with no data and regions with no data, however all the super-regions have at least one country contributing observed data.*
- *ii) For countries/regions with observed data, the sparseness of the data and/or size of the study sample are also reflected in the credible intervals of the estimate, which reflects the size of the sample of the studies included and the studies included contributing for that region.*
- *iii) We use a Bayesian model for exactly this reason, the sparseness of the data, and the prior distribution chosen is a key feature of the model in this context (something we have already stated). As far as we know, this approach is currently considered the gold standard in scenarios like this, and we kindly remind the reviewer that the model will be updated (please see further below).*

*We have previously indicated countries with no observed data in column "***Country with missing data***", eTable 6-11, Supplementary Material 2, and acknowledged in the manuscript, page 12, lines 14-18, that:*

> "Sixth, due to the complete lack of information for several countries and regions, in some circumstances it was only possible to use the estimate of the wider regions or super-regions for countries they were nested in. For this reason, estimates from countries with no data might be helpful in guiding policymakers, healthcare practitioners and patients but need to be interpreted with caution".

And added to the previous paragraph (page 12, lines 18-20):

> *"*In particular, estimates for countries with no observed data will likely be either an underestimate or an overestimate of the true value given they have been extrapolated from the regions or super-regions they are nested in.*"*

We also want to reassure the reviewer and the readers that one of the aims of the current study was to develop a methodological framework in order to estimate the prevalence/epidemiology of psoriasis worldwide. This is supposed to be the first iteration of a long-term project. The statistical model will be updated regularly by identifying new data and in future editions it will provide additional and more accurate estimates of the global

epidemiology of psoriasis. Therefore, the estimates will become more and more accurate through the time.
We have also added the following sentence in the Discussion, page 12, lines 20-23:

> *"*Nevertheless, given the long-term plan of the Global Psoriasis Atlas, future iterations of the model will include new data which will lead to more accurate estimates of the prevalence of psoriasis in particular in countries with no observed data*".*

2. The pattern of missing data is very uneven, with some regions and super regions having far fewer (if any) studies contributing to estimates than others. Where no data are available, my understanding is that the model estimates will entirely be based on "borrowed evidence" from higher levels in the model, such that the average region value is used for the country estimate when there is no country specific data, that the average super region value is used for countries when there are no data within a region, and that the global average is used when there are no data from within a super region. Looking at the results, particularly in the Figures, the data for the regions is presented is the same way regardless of whether it is based largely on data observed in the countries, or by "borrowing" from other regions. I believe that it is very important for readers to be able to distinguish observed results from extrapolated results. Could the authors look to augment the evidence presented in the Figures with a summary of the amount of the evidence is based on observed studies and how much is from borrowing strength? I do not know what graphical device or statistic could be used, but it is really important to indicate the real data from the extrapolated data, and the level at which the extrapolation is made.

The authors also need to be clear that the region and super region estimates where data are sparse are limited in that they are the values based on the small number of studies of studies in the same grouping. It is important that the reader is not mislead into thinking that they are based on regional or super regional level studies+.

A: As the reviewer has highlighted, for countries with no data available, the model estimates will be based on "borrowed evidence" from higher levels. However, i) the hierarchical model implies that country estimates are informed both by study data from the same country, if available, and by study data from other countries (within the same region); region estimates are informed both by study date from the same region, if available, and by study data from other regions (within the same super-region). Therefore, the higher-level units are not simply an average value, (eg average region value). As mentioned previously, as far as we know, this approach is considered the gold standard when data are sparse and heterogeneous. ii) We also want to clarify that while there are countries and regions with no observed data, all the seven super-regions in the geographical groups contain at least one country with observed data. We have now added this in the manuscript, page 7 lines 15-19:

> "In particular, for countries with no data available, the model estimates will be based on "borrowed evidence" from higher levels, such that the region estimate value is used for

the country estimate when there is no country specific data, that the super region estimate value is used for countries when there are no data within a region."

As the reviewer has pointed out, it would be very interesting to know the strength of the evidence based on observed studies and how much is from borrowing evidence from other countries/regions. However, we are not familiar with a method that would report something like this, in a Bayesian or even a frequentist context.
We also agree the readers need to be clear about what estimates are obtained from observed data and what estimates are obtained from higher levels. We have provided this information in eTables 6-11 that indicates whether a country had observed data or not ("Country with missing data"). We have now included an additional column, which indicates whether the estimates have been borrowed from the region or the super-regions. Finally, we have included to the main manuscript an additional Figure (Figure 1) which shows the number of studies for countries with observed data and have added to the main manuscript, page 7, lines 21-22:

"Information on countries with observed or missing data is reported in eTables 6-11, Supplementary Material 2".

We have now clarified in the main manuscript, page 7, lines 19-21, that the region and super region estimates are always based on studies coming at the country level and they are not regional or super-regional level studies.

"It is also noteworthy that the region and super-region estimates are always obtained from studies identified at the country level; therefore there are no regional or super-regional level studies".

3. **I cannot get all the figures in the text to match with those in the tables. For example Page 10 includes the statement:**

"Psoriasis occurred more frequently in adults than in children. In children, the prevalence of psoriasis varied between 0.02% (0.01% to 0.04%) in East Asia table states to 0.22% (0.06% to 0.81%) in Australasia and 0.21% (0.11% to 0.41%) in Western Europe (Figure 1)"

Several of the figures here do not agree with those in eTable 7. For example, the estimate of East Asia is 0.03 (0.01 to 0.09). Is there an explanation or are these errors? Similarly I cannot locate all the figures in the next statement:

"In adults, the disease varied between 0.14% (0.05% to 0.40%) in East Asia to 1.99% (0.64% to 6.60%) in Australasia. Other regions with an occurrence of the diseases above 1% were Western Europe 1.92% (1.07% to 3.46%); Central Europe 1.83% (0.62% to 5.32%); high-income North America 1.50% (0.63% to 3.60%); high-income Southern-Latin America 1.10% (0.36% to 2.96%); Figure 2."

*A: We are grateful to the reviewer for highlighting this issue and we apologise for this mistake. Unfortunately, in the previous submission we included the wrong Table in the Supplementary Material, such that the table reporting the "*Crude lifetime prevalence (physician or dermatologist diagnosed) of psoriasis in children" *(now eTable 6, Supplementary Material 2) was a duplicate of the table reporting the "Crude lifetime prevalence (self-reported diagnosis) of psoriasis in children" (now eTable 9, Supplementary Material 2). We have now included the correct table. We also want to highlight that the text in the manuscript is correct and matched with the estimate in eTable 6, Supplementary Material 2.*
The document containing the Supplementary Material is very long and, as previously mentioned, we have now split it into two documents to facilitate the reader to navigate through all the tables.

Similarly, there was a missing row for the region "High-income Southern Latin America" in the now eTable 7 which has been added.

We have also edited the category and used different colours in Figure 5 to highlight the difference in the prevalence of psoriasis between countries.

All the numbers reported in the manuscript, tables and figures now match.

4.There are also issues in the text concerning being clear where the reported values are from observed data and where it is based on extrapolation from the model. For example, in the statement immediately above, I do not believe that there were any studies in Southern-Latin America which contributed data, so this is an extrapolated value, whereas other examples are generated from the data. It would be helpful to have some way of separating out the observed data results from those which are based upon the model.

A: We understand the reviewer's concern. As explained in previous points, there is a column ("Country with missing data") in eTables 6-11 which indicates countries with missing data. We have also added a column to indicate whether the estimate of the country with no observed data comes from a higher-level region or super-region.
Furthermore, in the main manuscript, we have added a footnote to Figures 2-4 indicating region estimates which have been fully extrapolated and those that are not (i.e. the estimation used observed data as well).

5. The regions and super regions are not described in the main paper at the moment. When I found them in the additional document I was rather surprised by their structure, as they are largely based around income brackets and appear to vary between the different population groups and do not fit with geography as I recalled it (e.g. Argentina being classified as a North American country in eTable 8). It is very important for this structure to be described as it determines the way in which estimates for missing data are made and the rationale for this grouping to be explained and justified. It will have by itself created the income related pattern that has been commented on by others. If a different structure were used, estimates for countries with no data will

change according to the observations of other countries within the super region and region. The dependence of the findings upon both the categorisation of countries into regions and super regions definitely requires greater discussion and acknowledgement as a limitation in the discussion.

A: We apologies for the error in eTable 8 (now eTable 7). There was a missing row corresponding to the region "High-income Southern-Latin America" which belongs to the super-region "High-income" countries. eTable 7 has now been revised (Supplementary Material 2).
We would like to highlight that the income related pattern has also been found in previous studies investigating morbidity and mortality of skin diseases[1] and the relationship between socio-economic status and skin diseases in Europe[2].
In order to explain the structure and hierarchy of the countries, regions and super-regions further, we have now added to the main manuscript:

- *Table 4, which was previously included in the Supplementary material as eTable 6, with an explanation of the hierarchical structure, main manuscript, page 29.*

- *A sentence in the main manuscript, page 7, lines 4-5, which says:*

  "Countries were mapped according to the GBD classification with 189 countries nested in 21 regions, and regions nested in 7 super-regions, which hierarchy mainly follows geography and income."

- *A sentence to the Discussion, under "Limitations of the study", page 12, lines 10-14, which says:*

  "Fifth, the grouping and hierarchy of the countries within region and super-region, which aligns with the Global Burden of Disease methodology, is mainly based on geography and income. Therefore, it might have influenced the income related patterns highlighted in the findings."

- *A sentence to the Discussion, under "Comparison with other studies", page 13, lines 13-15, which says:*

  "ii) the structure of the hierarchical model used was mainly based on geography and income, which might have exacerbated the income related patterns of the findings."

1.Karimkhani C, Dellavalle RP, Coffeng LE, et al. Global Skin Disease Morbidity and Mortality: An Update From the Global Burden of Disease Study 2013Global Skin Disease Morbidity and MortalityGlobal Skin Disease Morbidity and Mortality. *JAMA Dermatology* 2017;153(5):406-12.
2. Ofenloch R, Schuttelaar M, Svensson Å, et al. Socioeconomic Status and the Prevalence of Skin and Atopic Diseases in Five European Countries. *Acta Derm Venereol* 2019;99:309–14.

6. I have also observed some discrepancies between the tables in the Appendix. For example, eTable 3 lists the countries which provide data on prevalence in children. There are no datasets listed here from sub-Saharan African countries. However, in eTable 7 there is a report that data are available for Tanzania which contradicts this eTable 3.

A: Thanks to the reviewer for highlighting this point, however we could not see any discrepancies between the tables. ETable 3 refers to the studies identified by the systematic review, whereas eTable 6 refers to the results produced by the statistical model. In the statistical model, we included a covariate (fixed effect) for "age strata" which was categorised in children, adults and the overall population, according to whether studies (from the systematic review) reported the prevalence on children, adults or the overall population. Then we used the regression coefficient (slope) of the covariate "age strata" to estimate the prevalence of psoriasis for each country (with observed data) within each age-stratum.

Furthermore, in order to facilitate access and readability of the supplementary information, the Supplementary Material has now been split into two documents:

- *Supplementary Material 1 (GPA_Paper_Supplementary_1_SystReview_BMJ_V3.docx), which includes all the information relative to the systematic review (stage I);*
- *and Supplementary Material 2, which includes all the information relative to the statistical model and meta-regression analysis.*