**SUPPORTING INFORMATION**

The impact of emerging *Plasmodium knowlesi* on accurate diagnosis by light microscopy: a systematic review and modelling analysis

John H. Huber, Margaret Elliot, Cristian Koepfli, T. Alex Perkins

Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN, USA

Corresponding Author: [huber.j.h@wustl.edu](mailto:huber.j.h@wustl.edu) (JHH)

**Supplemental Methods**

***Systematic Review***

Our search of MEDLINE and Web of Science returned 275 studies published before April 30th, 2021. After removing duplicates, we screened the abstracts of 176 studies. Of these, 101 full-text articles were then assessed for eligibility. Of the 75 studies that were excluded upon abstract screening, 50 were not primary research article and were instead case reports (n = 28), reviews (n = 14), conference abstracts (n = 3), editorials (n = 2), book chapters (n = 1), clinical guidelines n = 1), or study protocols (n = 1). The remaining studies excluded upon abstract screening did not compare diagnostic methods (n = 13), did not include *P. knowlesi* (n = 5), were conducted in a non-endemic location (n = 3), did not include clinical samples (n = 2), did not use a PCR protocol that targeted all five *Plasmodium* spp. (n = 1), or was a supplementary table to previously identified study (n = 1).

Of the 101 full-text articles that were assessed for eligibility, 12 met the inclusion criteria. The remaining 89 full-text articles were excluded for the following reasons: no comparison of diagnostic methods (n = 20); study conducted in non-endemic region (n = 12); no two-way table provided (n = 11); no presence of *P. knowlesi* (n = 10); review article (n = 8); no clinical samples (n = 6); case report (n = 4); conference abstract (n = 4); full-text unavailable (n = 4); PCR protocol could not target all five *Plasmodium* spp. (n = 3); *Plasmodium* spp. undefined (n = 2); overlapping data set with other identified publication (n = 2); no co-circulation of *Plasmodium* spp. parasites (n = 1); study not in English (n = 1); study conducted prior to *P. knowlesi* being recognized as public health threat (n = 1).

***Prisma Checklist***

| **Section and Topic** | **Item #** | **Checklist item** | **Location where item is reported** |
| --- | --- | --- | --- |
| **TITLE** | | |  |
| Title | 1 | Identify the report as a systematic review. | 1 |
| **ABSTRACT** | | |  |
| Abstract | 2 | See the PRISMA 2020 for Abstracts checklist. | 2 |
| **INTRODUCTION** | | |  |
| Rationale | 3 | Describe the rationale for the review in the context of existing knowledge. | 7 |
| Objectives | 4 | Provide an explicit statement of the objective(s) or question(s) the review addresses. | 7 |
| **METHODS** | | |  |
| Eligibility criteria | 5 | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses. | 8 |
| Information sources | 6 | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | 7 |
| Search strategy | 7 | Present the full search strategies for all databases, registers and websites, including any filters and limits used. | 7 |
| Selection process | 8 | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | 7-8, S2 |
| Data collection process | 9 | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | 8 |
| Data items | 10a | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | 8-9 |
| 10b | List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information. | 8-9 |
| Study risk of bias assessment | 11 | Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | 10, S13-S14 |
| Effect measures | 12 | Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results. | 10 |
| Synthesis methods | 13a | Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | 7-8, S2 |
| 13b | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions. | NA |
| 13c | Describe any methods used to tabulate or visually display results of individual studies and syntheses. | NA |
| 13d | Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used. | 8-10 |
| 13e | Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression). | 10, S13-S17 |
| 13f | Describe any sensitivity analyses conducted to assess robustness of the synthesized results. | 10, S13-S17 |
| Reporting bias assessment | 14 | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases). | S13-S17 |
| Certainty assessment | 15 | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome. | S7-S12 |
| **RESULTS** | | |  |
| Study selection | 16a | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram. | 10-11, S2 |
| 16b | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded. | S2 |
| Study characteristics | 17 | Cite each included study and present its characteristics. | 11, S9 |
| Risk of bias in studies | 18 | Present assessments of risk of bias for each included study. | S13-S14 |
| Results of individual studies | 19 | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots. | 10-15, S9-S12 |
| Results of syntheses | 20a | For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies. | S13-S14 |
| 20b | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | 10-15, S9-S12 |
| 20c | Present results of all investigations of possible causes of heterogeneity among study results. | 16-17, S13-S17 |
| 20d | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results. | S13-S17 |
| Reporting biases | 21 | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. | NA |
| Certainty of evidence | 22 | Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed. | S7-S12 |
| **DISCUSSION** | | |  |
| Discussion | 23a | Provide a general interpretation of the results in the context of other evidence. | 17-19 |
| 23b | Discuss any limitations of the evidence included in the review. | 19 |
| 23c | Discuss any limitations of the review processes used. | 19 |
| 23d | Discuss implications of the results for practice, policy, and future research. | 18 |
| **OTHER INFORMATION** | | |  |
| Registration and protocol | 24a | Provide registration information for the review, including register name and registration number, or state that the review was not registered. | NA |
| 24b | Indicate where the review protocol can be accessed, or state that a protocol was not prepared. | NA |
| 24c | Describe and explain any amendments to information provided at registration or in the protocol. | NA |
| Support | 25 | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review. | 10 |
| Competing interests | 26 | Declare any competing interests of review authors. | 10 |
| Availability of data, code and other materials | 27 | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review. | 21 |

***Accounting for Differences in Study Design***

Some studies in our systematic review reported samples as an “*P. knowlesi* / *P. malariae*” LM diagnosis, indicating that the microscopist identified *P. knowlesi* and/or *P. malariae* parasites in the sample but could not make a more precise mono-infection or co-infection diagnosis. We accounted for this in the latent class model by computing the union of all possible outcomes that could lead to a “*P. knowlesi* / *P. malariae*”, given a certain level of sensitivity and specificity of PCR diagnosis. This accounts for true and false positives for both *P. knowlesi* and *P. malariae* parasites.

***Hierarchical Model***

We fit our hierarchical latent class model to the studies identified in our systematic review to estimate the study-level sensitivities and specificities of LM as well as the hierarchical distributions for the five *Plasmodium* spp. Using eq. (1) in the main text, we computed the probability of observing a set of LM outcomes and PCR outcomes in study *l* across all *Plasmodium* spp. as

In eq. (S1), and are the vectors of LM sensitivities and specificities for all *Plasmodium* spp. in study *l*, and is the vector of prevalence for the *Plasmodium* spp. in study *l*. The individual probabilities in the product of eq. (S1) are computed using eq. (1).

Using eq. (1) and eq. (S1) and accounting for features of the study design, we computed a probability vector for each study *l*, where each element corresponds to the probability of observing a set of LM and PCR diagnostic outcomes across all *Plasmodium* spp. in that study. Given a vector containing the number of samples observed in study *l* for each corresponding set of LM and PCR diagnostic outcomes, we defined the likelihood of the model given the data as

We further defined parameters and and and , which comprised the hierarchical distributions for sensitivity and specificity for each *Plasmodium* spp. *k*. These distributions were defined on the logit scale, such that the study-level effects were computed as and for each *Plasmodium* spp. *k* and study *l*.

We used the No-U-Turn sampler provided in Stan to fit the model. The No-U-Turn sampler is an adaptive version of the Hamiltonian Monte Carlo algorithm that is capable of sampling from a high-dimensional posterior distribution. We assumed uniform prior distributions on [0,1] for the study-level sensitivities and specificities, uniform prior distributions on (-∞,∞) for the hierarchical means, and half-normal prior distributions with zero mean and standard deviation of 0.25 on the hierarchical standard deviations. We evaluated the sensitivity of our hierarchical parameter estimates to the assumed standard deviation of this half-normal distribution.

**Simulation Study**

To confirm that our inference framework could estimate each parameter with appropriate uncertainty, we performed a simulation study. We simulated 200 synthetic data sets, each consisting of 12 studies that matched the study designs of the 12 studies identified in our systematic review. For each synthetic data set, we sampled parameter sets from the respective prior distributions and simulated the data by using the likelihood of our model as a data-generating process. We then ran 4 chains with 2,000 samples each using the No-U-Turn Sampler in Stan with a warm-up period of 1,000 samples to obtain a posterior distribution of 4,000 samples for each synthetic data set. We assessed the performance of our inference framework by visually comparing the true values for each parameter to the posterior estimates that we obtained and by computing the coverage probabilities (i.e., the proportion of synthetic data sets for which the true value of each parameter is within the posterior credible interval). When calculated using the 95% credible interval, the coverage probability should be close to 0.95.

The results of our simulation study revealed that our inference framework could accurately estimate the parameters of our model (Fig. S1). We accurately estimated the means of the hierarchical distributions for sensitivity (Fig. S1A) and specificity (Fig. S1C), and the ranges of the coverage probabilities across the *Plasmodium* spp. was 0.89 – 0.95 and 0.93 – 0.97, respectively. There was greater uncertainty on our posterior estimates of hierarchical standard deviation for sensitivity (Fig. S1B) and specificity (Fig. S1D). Nevertheless, the range of the coverage probabilities across the *Plasmodium* spp. was 0.93 – 0.97 for the hierarchical standard deviation for sensitivity and 0.95 – 0.97 for the hierarchical standard deviation for specificity. In general, for study-level sensitivities, we observed good agreement between our posterior estimates that we obtained and the true values. There was greater uncertainty for our estimates for *P. knowlesi*, *P. malariae*, and *P. ovale* than our estimates for *P. falciparum* and *P. vivax* due to the lower observed PCR prevalence for these species. Nevertheless, the median (and range) coverage probabilities across the studies were 0.96 (0.92 – 0.98) for *P. falciparum*, 0.95 (0.93 – 0.98) for *P. vivax*, 0.96 (0.93 – 0.99) for *P. knowlesi*, 0.94 (0.91 – 0.96) for *P. malariae*, and 0.90 (0.86 – 0.91) for *P. ovale*. For study-level specificities, we also observed good agreement between our posterior estimates and the true values. The coverage probabilities were 0.95 (0.93 – 0.97) for *P. falciparum*, 0.95 (0.93 – 0.96) for *P. vivax*, 0.95 (0.92 – 0.97) for *P. knowlesi*, 0.96 (0.92 – 0.98) for *P. malariae*, and 0.95 (0.91 – 0.97) for *P. ovale*.

**Chart, line chart, scatter chart

Description automatically generated**

***Figure S1. Scatterplot comparison of true and inferred values from simulation study.*** *Comparisons of true and inferred posterior values are shown for the hierarchical (A) mean and (B) standard deviation of LM sensitivity, hierarchical (C) mean and (D) standard deviation of LM specificity, (E-I) study-level LM sensitivities, and (J-N) study-level LM specificities. In each plot, the point is the median posterior estimate, and the segment is the 95% credible interval. The color signifies the Plasmodium spp., with teal representing P. falciparum, orange representing P. vivax, purple representing P. knowlesi, pink representing P. malariae, and green representing P. ovale.*

**Systematic Review**

**Table S1. Characteristics of Included Studies.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Author** | **Publication Year** | **Location** | **Collection** | **Enrollment Criteria** | **Pk/Pm** | **Samples** |
| **Cooper *et al.*** | 2020 | Malaysia | Jan. 2015 – Dec. 2017 | PCR- and LM-Positive | Yes | 3,541 |
| **Han *et al.*** | 2017 | Myanmar | 2013 – 2015 | None | No | 90 |
| **Lubis *et al.*** | 2017 | Indonesia | Jan. 2015 – Jun. 2015 | None | No |  |
|  |  | *Batubara* |  |  |  | 1,270 |
|  |  | *Langkat* |  |  |  | 544 |
|  |  | *South Nias* |  |  |  | 1,917 |
| **Baum *et al.*** | 2016 | Thailand | Mar. 2012 – Jul. 2012 | None | No | 297 |
| **Chua *et al.*** | 2015 | Malaysia | 2008 – 2010 | None | No | 227 |
| **Yusof *et al.*** | 2014 | Malaysia | Sep. 2012 – Dec. 2013 | LM-Positive | No | 457 |
| **Zhou *et al.*** | 2014 | China | Jan. 2008 – Aug. 2012 | None | No | 560 |
| **Barber *et al.*** | 2013 | Malaysia | Sep. 2010 – Oct. 2011 | PCR- and LM-Positive | Yes | 303 |
| **Goh *et al.*** | 2013 | Malaysia | 2008 – 2011 | LM-Positive | No | 189 |
| **Barber *et al.*** | 2012 | Malaysia | Jan. 2009 – Nov. 2011 | LM-Positive | Yes | 461 |
| **Putaporntip *et al.*** | 2009 | Thailand | Oct. 2006 – Sep. 2007 | None | No | 1,874 |
| **Cox-Singh *et al*.** | 2008 | Malaysia | Mar. 2001 – Mar. 2006 | LM-Positive | No | 960 |

**Data Analysis**

***Summary of Results***

**Table S2. Posterior estimates and convergence statistics of group-level effects.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Posterior Estimate (95% CI)** | **Effective Sample Size** | **R-Hat** |
|  |  |  |  |
| *P. falciparum* | -0.332 (-1.22 – 0.579) | 954.8 | 1.005 |
| *P. vivax* | -0.394 (-1.26 – 0.469) | 1003 | 1.003 |
| *P. knowlesi* | -6.73 (-9.02 – -5.04) | 3917 | 1.001 |
| *P. malariae* | -2.50 (-3.60 – -1.27) | 3091 | 1.001 |
| *P. ovale* | -6.32 (-9.92 – -4.40) | 3996 | 1.000 |
|  |  |  |  |
| *P. falciparum* | 1.61 (1.26 – 2.09) | 2884 | 1.000 |
| *P. vivax* | 1.60 (1.26 – 2.04) | 2554 | 1.000 |
| *P. knowlesi* | 1.50 (0.996 – 2.12) | 5754 | 1.000 |
| *P. malariae* | 1.19 (0.43 – 1.91) | 2313 | 1.000 |
| *P. ovale* | 0.379 (0.0207 – 1.27) | 4418 | 0.9994 |
|  |  |  |  |
| *P. falciparum* | 4.20 (3.38 – 5.04) | 1152 | 1.007 |
| *P. vivax* | 4.38 (3.59 – 5.21) | 1224 | 1.001 |
| *P. knowlesi* | 8.03 (6.63 – 9.82) | 3662 | 1.000 |
| *P. malariae* | 5.02 (3.88 – 6.27) | 1310 | 1.005 |
| *P. ovale* | 11.10 (9.59 – 12.80) | 3849 | 1.000 |
|  |  |  |  |
| *P. falciparum* | 1.48 (1.13 – 1.93) | 2633 | 1.000 |
| *P. vivax* | 1.40 (1.06 – 1.86) | 3009 | 1.001 |
| *P. knowlesi* | 1.65 (1.18 – 2.27) | 5428 | 1.000 |
| *P. malariae* | 1.91 (1.47 – 2.44) | 4263 | 1.001 |
| *P. ovale* | 0.76 (0.0656 – 1.49) | 1846 | 1.000 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Posterior Estimate (95% CI)** | **Effective Sample Size** | **R-Hat** |
|  |  |  |  |
| Lubis *et al.* – Batubara | 36.8% (26.7 – 47.8%) | 4318 | 1.000 |
| Lubis *et al*. – Langkat | 29.0% (20.1 – 39.3%) | 4006 | 0.9994 |
| Lubis *et al*. – South Nias | 26.0% (20.7 – 31.9%) | 4223 | 0.9994 |
| Putaporntip *et al*. | 83.5% (80.7 – 86.0%) | 4158 | 1.001 |
| Han *et al.* | 90.6% (79.4 – 96.9%) | 5110 | 0.9995 |
| Baum *et al*. | 36.9% (20.3 – 56.5%) | 6167 | 1.000 |
| Cox-Singh *et al.* | 4.66% (2.36 – 8.62%) | 5323 | 0.9995 |
| Zhou *et al*. | 91.4% (86.4 – 95.1%) | 4220 | 0.9997 |
| Goh *et al.* | 42.0% (18.3 – 68.0%) | 4087 | 1.001 |
| Chua *et al*. | 92.7% (86.6 – 96.8%) | 4205 | 0.9997 |
| Yusof *et al*. | 11.0% (3.72 – 28.3%) | 2965 | 0.9998 |
| Cooper *et al*. | 0.962% (0.484 – 1.72%) | 5200 | 1.000 |
| Barber *et al*. | 50.3% (35.0 – 66.6%) | 4555 | 1.000 |
| Barber *et al.* | 42.6% (26.6 – 60.3%) | 4520 | 1.000 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 46.5% (37.9 – 55.2%) | 4351 | 1.000 |
| Lubis *et al*. – Langkat | 45.8% (35.7 – 56.7%) | 4100 | 1.000 |
| Lubis *et al*. – South Nias | 9.78% (6.94 – 13.4%) | 4173 | 1.000 |
| Putaporntip *et al*. | 83.5% (81.3 – 85.6%) | 3952 | 0.9994 |
| Han *et al.* | 80.4% (59.6 – 93.7%) | 6936 | 0.9997 |
| Baum *et al*. | 57.0% (38.8 – 74.0%) | 4967 | 0.9998 |
| Cox-Singh *et al.* | 5.16% (2.60 – 9.16%) | 5224 | 0.9997 |
| Zhou *et al*. | 90.7% (87.0 – 93.6%) | 4221 | 0.9993 |
| Goh *et al.* | 44.5% (19.0 – 72.7%) | 4412 | 1.001 |
| Chua *et al*. | 91.5% (84.6 – 96.2%) | 4081 | 0.9995 |
| Yusof *et al*. | 10.2% (3.47 – 23.8%) | 3137 | 0.9998 |
| Cooper *et al*. | 0.944% (0.470 – 1.74%) | 5465 | 1.001 |
| Barber *et al*. | 49.5% (38.5 – 60.6%) | 4503 | 0.9994 |
| Barber *et al.* | 28.2% (16.4 – 44.1%) | 4109 | 1.000 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 0.0824% (0.00224 – 0.817%) | 4488 | 0.9999 |
| Lubis *et al*. – Langkat | 0.104% (0.00239 – 1.57%) | 3559 | 1.000 |
| Lubis *et al*. – South Nias | 0.0619% (0.00208 – 0.522%) | 4491 | 1.000 |
| Putaporntip *et al*. | 0.116% (0.00225 – 2.40%) | 2554 | 1.000 |
| Han *et al.* | 0.126% (0.00242 – 3.17%) | 3079 | 0.9998 |
| Baum *et al*. | 0.117% (0.0021 – 3.64%) | 3541 | 0.9996 |
| Cox-Singh *et al.* | 0.0165% (0.000866 – 0.0812%) | 4559 | 1.000 |
| Zhou *et al*. | 0.117% (0.00204 – 2.79%) | 2572 | 1.001 |
| Goh *et al.* | 0.0849% (0.00198 – 0.928%) | 4605 | 0.9995 |
| Chua *et al*. | 0.102% (0.00251 – 1.46%) | 4054 | 1.000 |
| Yusof *et al*. | 7.38% (2.58 – 16.70%) | 2974 | 1.000 |
| Cooper *et al*. | 0.0324% (0.00104 – 0.186%) | 5038 | 0.9999 |
| Barber *et al*. | 68.2% (51.6 – 80.9%) | 7429 | 0.9994 |
| Barber *et al.* | 27.3% (15.3 – 41.0%) | 4935 | 0.9997 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 3.46% (0.317 – 14.3%) | 5931 | 0.9995 |
| Lubis *et al*. – Langkat | 10.3% (1.47 – 37.7%) | 6206 | 0.9998 |
| Lubis *et al*. – South Nias | 1.20% (0.141 – 4.63%) | 2656 | 1.000 |
| Putaporntip *et al*. | 7.43% (1.82 – 19.8%) | 5344 | 1.000 |
| Han *et al.* | 19.0% (2.43 – 82.2%) | 4572 | 0.9992 |
| Baum *et al*. | 18.5% (2.27 – 80.3%) | 3999 | 0.9995 |
| Cox-Singh *et al.* | 3.71% (0.909 – 11.3%) | 4785 | 1.000 |
| Zhou *et al*. | 4.15% (0.347 – 19.6%) | 5615 | 1.000 |
| Goh *et al.* | 20.9% (3.64 – 75.60%) | 3901 | 1.000 |
| Chua *et al*. | 31.7% (4.47 – 88.3%) | 4077 | 0.9998 |
| Yusof *et al*. | 4.19% (0.344 – 21.5%) | 4842 | 0.9998 |
| Cooper *et al*. | 0.0319% (0.00102 – 0.189%) | 4980 | 0.9993 |
| Barber *et al*. | 66.5% (13.4 – 98.6%) | 7362 | 0.9995 |
| Barber *et al.* | 57.8% (14.2 – 97.3%) | 6660 | 0.9994 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 0.177% (0.00419 – 1.65%) | 3921 | 1.000 |
| Lubis *et al*. – Langkat | 0.178% (0.00418 – 1.64%) | 4108 | 0.9995 |
| Lubis *et al*. – South Nias | 0.177% (0.00475 – 1.67%) | 4149 | 1.000 |
| Putaporntip *et al*. | 0.174% (0.00476 – 1.43%) | 4778 | 0.9998 |
| Han *et al.* | 0.177% (0.00463 – 1.74%) | 4174 | 0.9996 |
| Baum *et al*. | 0.179% (0.00377 – 1.68%) | 4243 | 0.9997 |
| Cox-Singh *et al.* | 0.168% (0.00444 – 1.14%) | 5030 | 0.9999 |
| Zhou *et al*. | 0.176% (0.00468 – 1.50%) | 5000 | 0.9995 |
| Goh *et al.* | 0.177% (0.00464 – 1.71%) | 4341 | 0.9996 |
| Chua *et al*. | 0.224% (0.00632 – 3.26%) | 4140 | 0.9994 |
| Yusof *et al*. | 0.179% (0.00451 – 1.47%) | 4701 | 0.9996 |
| Cooper *et al*. | 0.156% (0.00412 – 1.01%) | 5105 | 0.9999 |
| Barber *et al*. | 0.174% (0.00449 – 1.67%) | 3615 | 1.000 |
| Barber *et al.* | 0.177% (0.00454 – 1.59%) | 3948 | 0.9996 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Posterior Estimate (95% CI)** | **Effective Sample Size** | **R-Hat** |
|  |  |  |  |
| Lubis *et al.* – Batubara | 98.8% (98.2 – 99.3%) | 4383 | 1.001 |
| Lubis *et al*. – Langkat | 94.3% (92.0 – 96.2%) | 4080 | 1.000 |
| Lubis *et al*. – South Nias | 86.2% (84.5 – 87.8%) | 3990 | 0.9996 |
| Putaporntip *et al*. | 95.6% (94.3 – 96.7%) | 4151 | 1.000 |
| Han *et al.* | 97.1% (90.3 – 99.4%) | 7073 | 1.000 |
| Baum *et al*. | 99.8% (98.9 – 100%) | 6568 | 0.9994 |
| Cox-Singh *et al.* | 99.7% (99.4 – 99.8%) | 5361 | 0.9997 |
| Zhou *et al*. | 97.6% (95.9 – 98.8%) | 4422 | 0.9999 |
| Goh *et al.* | 97.1% (94.3 – 98.8%) | 4534 | 1.001 |
| Chua *et al*. | 98.0% (95.2 – 99.4%) | 6967 | 1.000 |
| Yusof *et al*. | 99.8% (99.5 – 99.9%) | 3607 | 0.9998 |
| Cooper *et al*. | 100% (100 – 100%) | 5319 | 1.000 |
| Barber *et al*. | 93.7% (91.2 – 95.6%) | 4344 | 0.9992 |
| Barber *et al.* | 92.7% (88.8 – 95.6%) | 4468 | 0.9997 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 98.1% (97.2 – 98.8%) | 4348 | 0.9998 |
| Lubis *et al*. – Langkat | 96.5% (94.6 – 97.9%) | 4860 | 0.9995 |
| Lubis *et al*. – South Nias | 93.8% (92.5 – 94.9%) | 4088 | 0.9996 |
| Putaporntip *et al*. | 96.5% (94.9 – 97.6%) | 3826 | 1.001 |
| Han *et al.* | 98.0% (93.4 – 99.6%) | 7385 | 0.9994 |
| Baum *et al*. | 99.8% (99.0 – 100%) | 5548 | 0.9996 |
| Cox-Singh *et al.* | 99.5% (99.1 – 99.8%) | 4934 | 0.9993 |
| Zhou *et al*. | 95.2% (92.3 – 97.3%) | 4324 | 0.9993 |
| Goh *et al.* | 98.5% (96.5 – 99.5%) | 5187 | 1.000 |
| Chua *et al*. | 98.7% (96.1 – 99.7%) | 5386 | 1.000 |
| Yusof *et al*. | 99.6% (99.0 – 99.9%) | 3301 | 1.001 |
| Cooper *et al*. | 100% (100 – 100%) | 5082 | 0.9999 |
| Barber *et al*. | 96.7% (94.9 – 98.0%) | 4101 | 0.9997 |
| Barber *et al.* | 96.5% (94.1 – 98.1%) | 4310 | 1.002 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 100% (99.9 – 100%) | 4745 | 0.9994 |
| Lubis *et al*. – Langkat | 100% (99.8 – 100%) | 3848 | 0.9997 |
| Lubis *et al*. – South Nias | 100% (99.9 – 100%) | 4942 | 0.9998 |
| Putaporntip *et al*. | 100% (99.9 – 100%) | 4737 | 0.9999 |
| Han *et al.* | 100% (99.5 – 100%) | 3858 | 1.000 |
| Baum *et al*. | 100% (99.7 – 100%) | 4708 | 0.9999 |
| Cox-Singh *et al.* | 100% (100 – 100%) | 5397 | 0.9996 |
| Zhou *et al*. | 100% (99.8 – 100%) | 3738 | 0.9992 |
| Goh *et al.* | 100% (99.8 – 100%) | 4368 | 0.9994 |
| Chua *et al*. | 100% (99.6 – 100%) | 3778 | 0.9998 |
| Yusof *et al*. | 99.6% (99.0 – 99.9%) | 3827 | 0.9997 |
| Cooper *et al*. | 95.4% (92.3 – 97.6%) | 5171 | 1.001 |
| Barber *et al*. | 100% (98.7 – 100%) | 2026 | 1.002 |
| Barber *et al.* | 100% (95.4 – 100%) | 966 | 1.000 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 100% (99.8 – 100%) | 6584 | 0.9994 |
| Lubis *et al*. – Langkat | 99.9% (99.5 – 100%) | 6029 | 0.9993 |
| Lubis *et al*. – South Nias | 100% (99.8 – 100%) | 5447 | 0.9993 |
| Putaporntip *et al*. | 99.9% (99.7 – 100%) | 5018 | 1.000 |
| Han *et al.* | 99.8% (98.1 – 100%) | 4800 | 0.9997 |
| Baum *et al*. | 99.9% (99.2 – 100%) | 4453 | 1.001 |
| Cox-Singh *et al.* | 98.0% (96.4 – 99.0%) | 5302 | 0.9997 |
| Zhou *et al*. | 99.9% (99.5 – 100%) | 5715 | 0.9998 |
| Goh *et al.* | 90.3% (84.4 – 95.4%) | 3733 | 1.001 |
| Chua *et al*. | 86.7% (81.8 – 90.7%) | 4103 | 1.001 |
| Yusof *et al*. | 98.1% (95.6 – 99.3%) | 3040 | 0.9999 |
| Cooper *et al*. | 98.7% (97.7 – 99.3%) | 5303 | 1.000 |
| Barber *et al*. | 59.9% (48.4 – 70.6%) | 3881 | 0.9996 |
| Barber *et al.* | 91.5% (86.8 – 96.7%) | 2244 | 1.001 |
|  |  |  |  |
| Lubis *et al.* – Batubara | 100% (100 – 100%) | 3588 | 0.9997 |
| Lubis *et al*. – Langkat | 100% (100 – 100%) | 4064 | 1.000 |
| Lubis *et al*. – South Nias | 100% (100 – 100%) | 3532 | 1.002 |
| Putaporntip *et al*. | 100% (100 – 100%) | 3645 | 0.9995 |
| Han *et al.* | 100% (100 – 100%) | 3705 | 1.000 |
| Baum *et al*. | 100% (100 – 100%) | 3262 | 1.000 |
| Cox-Singh *et al.* | 100% (100 – 100%) | 4457 | 1.000 |
| Zhou *et al*. | 100% (100 – 100%) | 3285 | 1.000 |
| Goh *et al.* | 100% (100 – 100%) | 3403 | 0.9996 |
| Chua *et al*. | 100% (100 – 100%) | 3704 | 1.000 |
| Yusof *et al*. | 100% (100 – 100%) | 4087 | 0.9991 |
| Cooper *et al*. | 100% (100 – 100%) | 3898 | 0.9996 |
| Barber *et al*. | 100% (100 – 100%) | 3965 | 1.000 |
| Barber *et al.* | 100% (100 – 100%) | 3959 | 0.9997 |

***Validation of Results***

***Diagram

Description automatically generated***

***Figure S2. Posterior predictive check of fitted model.*** *Comparison of the observed data and the posterior predictions are made for each of the studies included in model fitting. On the horizontal axis, each observation is a data point from a study, representing the number of samples that were observed with a given set of LM and PCR diagnostic outcomes across all Plasmodium spp. On the vertical axis, the point is the median posterior prediction, and the segment is the 95% prediction interval. The diagonal line in each plot is the one-to-one line, and the coverage probabilities and R2 are reported for each study. Both axes are on a base-10 logarithm scale.*

***Sensitivity Analyses***

*Sensitivity to the Included Studies*

To assess whether the inclusion of certain studies biased the results of our meta-analysis, we performed a sensitivity analysis for each study in which we re-fit the model while excluding that study. If the estimates of our hierarchical distribution means are robust to inclusion or exclusion of a given study, then that provides support that that study did not bias the parameter estimates that we obtained.

Our sensitivity analysis revealed that the hierarchical means are robust to the inclusion and exclusion of the studies (Fig. S3). For *P. falciparum*, *P. vivax*, and *P. malariae,* we observed variation in posterior median estimates of sensitivity and specificity for each study excluded. However, the uncertainty around these estimates was wide, and the 95% credible intervals overlapped with the estimates from the full analysis, indicating that our results were robust to the inclusion and exclusion of each study. For *P. knowlesi* and *P. ovale*, our estimates of sensitivity and specificity changed little with each study excluded.

Chart

Description automatically generated

***Figure S3. Sensitivity of hierarchical means to the included studies.*** *The hierarchical means for (A) sensitivity and (B) specificity are reported for the study excluded during model fitting. The colors denote the Plasmodium spp. considered, with teal representing P. falciparum, orange representing P. vivax, purple representing P. knowlesi, pink representing P. malariae, and green representing P. ovale. Each point is the posterior median estimate, and the vertical segment is the 95% credible interval.*

*Sensitivity to the Assumption of PCR as the Gold Standard*

In the primary analysis, we treated PCR as the gold standard and thus assumed that and for all *Plasmodium* spp. *k* and studies *l*. To evaluate how robust our inferences were to this assumption, we performed a sensitivity analysis in which we alternative assumed that and for all *Plasmodium* spp. *k* and studies *l*. Accordingly, we computed the true parasite prevalence of *Plasmodium* spp. *k* in study *l* as

In eq. (S3), is the observed PCR prevalence when assuming PCR as the gold standard. The value of the specificity was chosen to ensure that for all *Plasmodium* spp. *k* and studies *l*.

We found that our posterior estimates of the species-level sensitivities and specificities were mostly robust to the assumed values of PCR diagnostic performance (Fig. S4). Assuming an imperfect PCR method, our species-level estimates of sensitivity for *P. falciparum*, *P. vivax*, and *P. ovale* increased slightly, though the credible intervals were wide and generally overlapped with the corresponding estimates obtained under the assumption of PCR as the gold standard. Our estimates of species-level specificity changed very little with the assumed values of PCR diagnostic performance.

Chart

Description automatically generated

**Figure S4. Sensitivity of LM diagnostic performance to the assumption of PCR as the gold standard.** The posterior means for LM sensitivity (circles) and specificity (triangles) are shown when assuming PCR as the gold standard (filled points) and when assuming an imperfect PCR method (unfilled points) for each of the five *Plasmodium* spp. that causes human malaria. Each point is the posterior median estimate, and the vertical segment is the 95% credible interval.

*Sensitivity to the Prior Distribution*

In our inference framework, we placed a normal prior distribution with mean of zero and standard deviation of 0.25 on hierarchical standard deviation parameters. To evaluate whether our inferences were sensitive to the assumed standard deviation of this prior distribution, we re-fit our model, assuming values of prior distribution standard deviation from 0.125 to 1.0 in increments of 0.125.

In our sensitivity analysis, the hierarchical means for both sensitivity and specificity were robust to the assumed prior standard deviation (Fig. S5A & B). As the assumed standard deviation of the prior distribution increased, the posterior estimates for the hierarchical means increased only slightly. The width of the 95% credible interval increased with increasing prior standard deviation. That the hierarchical means for both sensitivity and specificity across the *Plasmodium* spp. were robust to the assumed prior distribution standard deviation further supports the conclusions reached in the primary analysis.

The estimates for the hierarchical standard deviations increased as the standard deviation on the prior distribution increased. This effect was more pronounced, given that the prior distribution was directly placed on these parameters. Although the estimate of the hierarchical standard deviation is sensitive to the prior assumption, the conclusions that we reached in the primary analysis do not depend strongly upon these parameters, so this should not affect the conclusions of our study.

*Chart, scatter chart

Description automatically generated*

***Figure S5. Sensitivity of hierarchical parameters to prior distribution assumption.*** *The sensitivities of the hierarchical mean for (A) sensitivity and (B) specificity and the hierarchical standard deviations of (C) sensitivity and (D) specificity are shown for each assumed prior standard deviation describing the hierarchical standard deviation parameters. In each plot, the point is the posterior median, and the vertical segment is the 95% credible interval. The color denotes the Plasmodium spp., with teal representing P. falciparum, orange representing P. vivax, purple representing P. knowlesi, pink representing P. malariae, and green representing P. ovale.*