The Impact of Emerging \textit{Plasmodium knowlesi} on Accurate Diagnosis by Light Microscopy: A Systematic Review and Modeling Analysis

John H. Huber,* Margaret Elliott, Cristian Koepfli, and T. Alex Perkins
Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, Indiana

Abstract. The five major \textit{Plasmodium} spp. that cause human malaria appear similar under light microscopy, which raises the possibility that misdiagnosis could routinely occur in clinical settings. Assessing the extent of misdiagnosis is of particular importance for monitoring \textit{P. knowlesi}, which co-circulates with the other \textit{Plasmodium} spp. We performed a systematic review and meta-analysis of studies comparing the performance of microscopy and polymerase chain reaction (PCR) for diagnosing malaria in settings with co-circulation of the five \textit{Plasmodium} spp. We assessed the extent to which co-circulation of \textit{Plasmodium} parasites affects diagnostic outcomes. We fit a Bayesian hierarchical latent class model to estimate variation in microscopy sensitivity and specificity measured against PCR as the gold standard. Mean sensitivity of microscopy was low, yet highly variable across \textit{Plasmodium} spp., ranging from 65.7% (95% confidence interval: 48.1–80.3%) for \textit{P. falciparum} to 0.525% (95% confidence interval 0.0210–3.111%) for \textit{P. ovale}. Observed PCR prevalence was positively correlated with estimated microscopic sensitivity and negatively correlated with estimated microscopic specificity, though the strength of the associations varied by species. Our analysis suggests that cocirculation of \textit{Plasmodium} spp. undermines the accuracy of microscopy. Sensitivity was considerably lower for \textit{P. knowlesi}, \textit{P. malariae}, and \textit{P. ovale}. The negative association between specificity and prevalence imply that less frequently encountered species may be misdiagnosed as more frequently encountered species. Together, these results suggest that the burden of \textit{P. knowlesi}, \textit{P. malariae}, and \textit{P. ovale} may be underappreciated in a clinical setting.

INTRODUCTION

\textit{Plasmodium falciparum} and \textit{Plasmodium vivax} are the predominant parasites responsible for human malaria worldwide, contributing nearly 241 million clinical cases in 2020.\cite{WHO2020} Given their contribution to the global burden of malaria, \textit{P. falciparum} and \textit{P. vivax} remain the primary focus of public health efforts aimed at achieving malaria control, elimination, and eradication.\cite{WHO2018} However, three additional major \textit{Plasmodium} species—\textit{P. knowlesi}, \textit{P. malariae}, and \textit{P. ovale}—also cause clinical episodes of malaria, although their overall burden remains uncertain due in part to routine misdiagnosis in a clinical setting.\cite{WHO2020}

Diagnosis of clinical malaria is routinely performed using light microscopy (LM). In addition to challenges presented by low or submicroscopic parasite densities,\cite{WHO2021} accurate speciation of malaria parasites using LM depends on identification of morphological features characteristic of the \textit{Plasmodium} spp. causing the infection. However, morphological similarities across the five \textit{Plasmodium} spp., such as the resemblance of early blood-stage \textit{P. knowlesi} parasites to \textit{P. falciparum} parasites and late blood-stage \textit{P. knowlesi} parasites to \textit{P. malariae} parasites,\cite{WHO2021} undermine species-level identification of \textit{Plasmodium} infections.\cite{WHO2021}

In settings with co-circulation of the \textit{Plasmodium} spp., routine misdiagnosis by light microscopy has consequences for clinical decision-making and malaria control. The choice of therapeutic is specific to the \textit{Plasmodium} spp. causing the infection,\cite{WHO2021} so misdiagnosis may lead to ineffective treatment of blood-stage parasites for potentially life-threatening \textit{Plasmodium} spp. infections as well as ineffective or no treatment of liver-stage parasites for \textit{P. vivax} and \textit{P. ovale} infections. Additionally, misdiagnosis may present challenges for monitoring \textit{Plasmodium} spp. transmission, particularly for the emerging zoonosis \textit{P. knowlesi}. Currently, the extent of human-to-human transmission of \textit{P. knowlesi} is unknown,\cite{WHO2021} and misdiagnosis could cause transmission clusters to appear smaller than their actual size, leading to an underestimate of transmission.\cite{WHO2021}

Consequently, measuring the extent of misdiagnosis by light microscopy is of importance to public health, and past studies have directly compared the diagnostic performance of light microscopy to polymerase chain reaction (PCR) in settings with co-circulation of \textit{Plasmodium} spp. parasites.\cite{WHO2021} However, these studies varied in their criteria for enrollment and other study design features, as well as in the epidemiological settings in which they were conducted. This makes direct comparison of their estimates of diagnostic performance challenging.

To synthesize across these studies, we performed a systematic review and meta-analysis of the diagnostic performance of light microscopy in settings with co-circulation of all five \textit{Plasmodium} spp. We identified studies that compared the diagnostic performance of light microscopy to PCR and fit a hierarchical Bayesian latent class model to estimate study- and species-level sensitivities and specificities of light microscopy. Finally, to assess whether the perceived prevalences of the five species might affect diagnostic outcomes of light microscopy, we explored the relationship between the estimates of sensitivity and specificity that we obtained and the observed PCR prevalences in each study.

METHODS

Search strategy and selection criteria. In this systematic review and meta-analysis, JHH searched MEDLINE and Web of Science for studies published in English before April 30, 2021, that compared the diagnostic performance of LM and PCR for \textit{Plasmodium} spp. infections. We identified studies using the following search strategy: (\textit{Plasmodium} OR malaria) AND knowlesi AND (microscopy OR microscopic OR blood film OR thick film OR thin film) AND (PCR or “polymerase chain reaction”).
Studies met the criteria for inclusion if they compared the performance of LM and PCR on clinical samples collected in areas with endemic transmission of all five Plasmodium spp. that naturally infect humans (i.e., P. falciparum, P. vivax, P. knowlesi, P. malariae, and P. ovale). Furthermore, studies were only included if they used a PCR protocol that was capable of amplifying DNA from all five Plasmodium spp. Before 2004, the public health burden of P. knowlesi was underappreciated, so restricting our search to studies published between 2004 and 2021 ensured that the diagnostic outcomes were representative of the current clinical and diagnostic landscape.

The data were extracted from each study that met the criteria for inclusion. Because the data in each study were anonymized and reported in aggregate, ethical approvals were not required according to a determination from the University of Notre Dame Institutional Review Board that this study did not constitute human subjects research.

**Data analysis.** Overview. To model the diagnostic performance of LM and PCR, we constructed a Bayesian hierarchical latent class model of clinical diagnosis. In this model, the probability \( P_{ij} \) of obtaining LM outcome \( i \in \{1, 0\} \) and PCR outcome \( j \in \{1, 0\} \) for Plasmodium spp. in study \( k \) is

\[
\begin{align*}
P_{ij} &= \frac{se_{LM}^{(ij)} \cdot sp_{LM}^{(ij)} \cdot \left[ \left( 1 - sp_{PCR}^{(ij)} \right) \left( 1 - sp_{LM}^{(ij)} \right) \right] + \left[ \left( 1 - sp_{LM}^{(ij)} \right) \left( 1 - sp_{PCR}^{(ij)} \right) \right] \cdot \left( 1 - \theta_{LM}^{(ij)} \right), \quad i = 1, j = 1 \\
&+ \frac{se_{LM}^{(ij)} \cdot \left( 1 - sp_{PCR}^{(ij)} \right) \cdot \left( 1 - sp_{LM}^{(ij)} \right) \cdot \theta_{LM}^{(ij)}, \quad i = 1, j = 0 \\
&+ \frac{se_{LM}^{(ij)} \cdot \left( 1 - sp_{PCR}^{(ij)} \right) \cdot sp_{LM}^{(ij)} \cdot \left( 1 - \theta_{LM}^{(ij)} \right), \quad i = 0, j = 1 \\
&+ \left( 1 - se_{LM}^{(ij)} \right) \cdot \left( 1 - sp_{PCR}^{(ij)} \right) \cdot \left( 1 - sp_{LM}^{(ij)} \right), \quad i = 0, j = 0
\end{align*}
\]

In Eq. (1), \( se_{LM}^{(ij)} \) and \( sp_{LM}^{(ij)} \) are the sensitivity and specificity of LM for Plasmodium spp. \( k \) in study \( l \), \( se_{PCR}^{(ij)} \) and \( sp_{PCR}^{(ij)} \) are the sensitivity and specificity of PCR for Plasmodium spp. \( k \) in study \( l \), and \( \theta_{LM}^{(ij)} \) is the prevalence of Plasmodium spp. \( k \) among participants in study \( l \). In many cases, \( \theta_{LM}^{(ij)} \) is not equivalent to community prevalence, due to characteristics of the study participants that are not representative of the surrounding community.

A key assumption of the latent class model is that diagnostic outcomes are conditionally independent. That is, by assuming that the sensitivity and specificity of a given diagnostic method is unique for each Plasmodium spp., the outcomes we observe for one parasite species using a given diagnostic method is modeled as independent of the outcomes observed for all other parasite species and diagnostic methods. This implies that the probability of a set of diagnostic outcomes observed across all five Plasmodium spp. is equal to the product of the probabilities of the diagnostic outcome for each Plasmodium spp. Additionally, the probability of acquiring a mixed infection is equal to the product of the prevalence of each Plasmodium spp. present within that mixed infection. Nevertheless, the assumption of conditional independence still allows for interactions between Plasmodium spp. that arise in the context of misdiagnosis. For example, if a P. knowlesi infection is misdiagnosed as a P. malariae infection using LM, then the sensitivity of LM for P. knowlesi and the specificity of LM for P. malariae are penalized because both a false-negative P. knowlesi diagnosis and a false-positive P. malariae diagnosis occurred.

Accounting for differences in study design. Studies identified in the systematic review varied by enrollment and diagnostic criteria. Where enrollment in a study required a positive Plasmodium spp. diagnosis by LM or PCR, we normalized the probabilities of Eq. (1), such that they were consistent with the subset of possible diagnostic outcomes that could be observed with that study design.

Because P. knowlesi and P. malariae share morphological similarities, studies may diagnose samples as "P. knowlesi/P. malariae" with LM, 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. To accommodate this in our model, we defined parameters \( se_{LM}^{(i/0, j/0)}, se_{LM}^{(0/1, j/0)} \) as the sensitivities of LM to diagnose samples in study \( l \) as "P. knowlesi/P. malariae," given that the samples contained P. knowlesi and P. malariae parasites, respectively.

Inference. Using PCR as the gold standard, we fit a hierarchical latent class model of clinical diagnosis to the studies identified in our systematic review to estimate study-level and species-level variation in LM sensitivity and specificity for each Plasmodium spp. We fit our model in a Bayesian framework using Stan by running four independent chains of 2,000 samples, each with a warm-up period of 1,000 samples. We tested for convergence with the Gelman–Rubin statistic and pooled the independent chains for a posterior distribution of 4,000 samples. Supplementary analyses were performed to evaluate the sensitivity of our posterior estimates to the inclusion of each study and to the assumption of PCR as the gold standard. See the supplemental information for more details.

**RESULTS**

Of the 176 unique studies identified using MEDLINE and Web of Science, 10 studies met the inclusion criteria, and their data comparing LM and PCR diagnostic performance were extracted (Figure 1). The studies enrolled febrile patients presenting in health clinics in Malaysia \((N = 6)\), Thailand \((N = 2)\), Indonesia \((N = 1)\), Myanmar \((N = 1)\), and China \((N = 1)\). Among the 7,999 total patients enrolled across the 12 studies, the pooled percentage positive by PCR was 19% for P. falciparum, 25% for P. vivax, 52% for P. knowlesi, 1.1% for P. malariae, and 0.48% for P. ovale (Table 1). The studies varied widely in the percentage positive by PCR across Plasmodium spp., encompassing a wide range of epidemiological settings and study designs (i.e., criteria for patient enrollment) in which we estimated LM diagnostic performance.

After validating on 200 simulated datasets (Supplemental Figure 1), we applied our hierarchical latent class model to the data from the 10 studies identified in the systematic review. Our fitted hierarchical latent class model captured the variation in the observed data (Supplemental Figure 6) and revealed differences in LM diagnostic performance across Plasmodium spp. (Figure 2). Specifically, we estimated that the species-level mean sensitivity of LM was 65.7% (95% confidence interval [CI]: 48.1–90.3%) for P. falciparum, 62.7% (95% CI: 46.0–77.2%) for P. vivax, and 52.5% (95% CI: 0.0210–3.11%) for P. ovale. It was not possible to estimate group-level mean sensitivities for the pooled P. knowlesi and P. malariae LM diagnosis due to identifiability concerns, yet
the mean of study-level posterior median estimates were 44.9% for *P. knowlesi* parasites and 45.3% for *P. malariae* parasites. Species-level mean specificity of LM was consistently high across *Plasmodium* spp. and was estimated at 98.2% (95% CI: 96.6–99.1%) for *P. falciparum*, 98.6% (95% CI: 97.3–99.3%) for *P. vivax*, and 100% (95% CI: 100%–100%) for *P. ovale*. For the pooled *P. knowlesi* and *P. malariae* LM diagnosis, the mean of study-level posterior median estimates of specificity were 95.6% for *P. knowlesi* parasites and 95.9% for *P. malariae* parasites. These results were robust to the assumption of PCR as the gold standard (Supplemental Figure 9).

The variance in the hierarchical distributions of LM diagnostic performance in Figure 2 revealed the extent of variation across studies, a feature that depended on both the diagnostic measure and the *Plasmodium* spp. considered (Figure 3). In general, for a given *Plasmodium* spp., we estimated greater variation across studies in LM sensitivity than specificity. For example, for *P. falciparum*, we estimated that logit-transformed standard deviation of LM sensitivity was 1.09 (0.886–1.36), suggesting that the range of sensitivities that fell within one standard deviation of the mean was 60.9% to 85.1%. By comparison, the estimated logit-transformed standard deviation of LM specificity for *P. falciparum* was 0.967 (0.753–1.23), corresponding to a range of specificities within one standard deviation of the mean of 95.4% to 99.3%. Across *Plasmodium* spp., study-level estimates of LM sensitivity and specificity for *P. falciparum*, *P. vivax*, *P. knowlesi*, and *P. malariae* exhibited greater variation than for *P. ovale*.

Because the estimates of LM diagnostic performance varied across the unique studies, we then assessed its relationship with observed PCR prevalence (Figure 4). We observed strong, positive correlations between PCR prevalence and LM sensitivity for *P. vivax* (Pearson correlation test: $r = 0.78; P = 0.0082$) and *P. falciparum* ($r = 0.62; P = 0.054$) and weaker, positive correlations for *P. knowlesi* ($r = 0.01; P = 0.97$) and *P. ovale* ($r = 0.06; P = 0.087$), although the latter two were not statistically significant associations. By contrast, for *P. malariae*, there was a negative correlation between PCR prevalence and LM sensitivity, although this estimate was also not statistically significant. Across *Plasmodium* spp., increased PCR prevalence of *P. knowlesi* was associated with reduced sensitivity of LM for *P. falciparum* ($r = -0.06; P = 0.042$) and *P. vivax* ($r = -0.76; P = 0.010$), highlighting a possible source of misdiagnosis. Additional negative associations were observed between PCR prevalence and LM sensitivity across *P. falciparum*, *P. knowlesi*, and *P. malariae—Plasmodium* spp. that are commonly misdiagnosed, although these quantities were not statistically significant.

Generally, for a given *Plasmodium* spp., PCR prevalence was negatively associated with LM specificity (Figure 4B). For *P. knowlesi*, this association was strong and statistically significant ($r = -0.79; P = 0.0061$), and additional weaker, yet statistically nonsignificant relationships were observed for *P. falciparum*, *P. vivax*, and *P. ovale*. The opposite relationship was observed for *P. malariae*, although this quantity was not statistically significant. Across *Plasmodium* spp.,
<table>
<thead>
<tr>
<th>Study Location</th>
<th>Year</th>
<th>Type of Samples</th>
<th>Light Microscopy (LM) Percentage Positive</th>
<th>PCR Percentage Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thailand</td>
<td>2006</td>
<td>Any Febrile</td>
<td>19 (1,540)</td>
<td>9.1 (27)</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2008-2011</td>
<td>38 (70)</td>
<td>0.53 (10)</td>
<td>22 (42)</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2009-2011</td>
<td>22 (54)</td>
<td>1.1 (4)</td>
<td>45 (96)</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2010-2012</td>
<td>22 (42)</td>
<td>0.86 (19)</td>
<td>12 (25)</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2012-2013</td>
<td>32 (64)</td>
<td>0.36 (8)</td>
<td>30 (62)</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2013-2017</td>
<td>40 (81)</td>
<td>1.22 (26)</td>
<td>50 (101)</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2015-2017</td>
<td>49 (96)</td>
<td>1.1 (24)</td>
<td>50 (101)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>7,999</td>
<td>52 (1,036)</td>
</tr>
</tbody>
</table>

In this systematic review and meta-analysis, we estimated study- and species-level variation in the malaria diagnostic performance of light microscopy in epidemiological settings with cocirculation of the five major *Plasmodium* spp. that cause human malaria. Our analysis revealed variable sensitivity of light microscopy to diagnose *Plasmodium* spp. infection. Sensitivity of light microscopy was estimated to be considerably lower for *P. knowlesi*, *P. malariae*, and *P. ovale* than for *P. falciparum* and *P. vivax*. This suggests that the burden of *P. knowlesi*, *P. malariae*, and *P. ovale* in a clinical setting may be appreciably underestimated, consistent with a recent study of *P. ovale* in health clinics in Kenya.³

Post hoc analyses revealed relationships between the diagnostic performance of light microscopy and the epidemiological setting. In general, we noted a positive association between the PCR prevalence of a *Plasmodium* spp. and the sensitivity of light microscopy for that *Plasmodium* spp. This supports a phenomenon whereby the more a microscopist encounters samples of a *Plasmodium* spp. infection in a clinical setting, the more likely they are to accurately identify the *Plasmodium* spp. causing that infection. By contrast, increased PCR prevalence of a *Plasmodium* spp. was associated with high false-positive probabilities for that *Plasmodium* spp. This tendency to misdiagnose malaria infections as what is most commonly encountered in that clinical setting could lead to ineffective monitoring of less commonly encountered *Plasmodium* spp., such as *P. knowlesi*, *P. malariae*, and *P. ovale*, and especially could present challenges in settings where a *Plasmodium* spp. is newly emerging.

The high probabilities of misdiagnosis estimated in this analysis could have important consequences for malaria treatment and control. Specifically, we observed an association between increased prevalence of *P. knowlesi* and reduced sensitivity of LM for *P. vivax*. This suggests that in areas with *P. knowlesi*, infections of *P. vivax* may not be correctly diagnosed and therefore may not receive radical cure therapy to clear hypnozoites and prevent relapse.²⁴ Additionally, increased prevalence of *P. malariae* was associated with reduced specificity of LM for *P. vivax*, indicating that many *P. malariae* infections may be misdiagnosed as *P. vivax*. This form of misdiagnosis could unnecessarily subject individuals to radical cure therapy, potentially increasing the risk of acute hemolysis among individuals deficient of the glucose-6-phosphate dehydrogenase enzyme.²⁵ The potential for misdiagnosis in the setting of cocirculation should thus be a consideration when guiding treatment recommendations, although this will likely depend strongly on local epidemiology.

Routine diagnosis of clinical malaria by LM may prevent effective monitoring of *P. knowlesi* transmission and burden. We estimated variable sensitivity of LM to diagnose *P. knowlesi* and *P. malariae* infections, given the morphological similarities of *P. falciparum*, *P. knowlesi*, and *P. malariae* parasites²⁵ as well as likely delays in familiarity of...
Figure 2. Group-level posterior estimates of light microscopy (LM) diagnostic performance. The group-level distributions of LM sensitivity (A, C, E) and LM specificity (B, D, F) are shown for Plasmodium falciparum (teal; A, B), Plasmodium vivax (orange; C, D), and Plasmodium ovale (green; E, F). Thick lines are the maximum a posteriori estimates, and thin lines are 25 posterior samples.
P. knowlesi among microscopists in the health systems. Although in practice P. knowlesi and P. malariae infections diagnosed by light microscopy are confirmed by PCR in many settings, P. knowlesi infections misdiagnosed as P. falciparum, for instance, would not be subject to confirmatory PCR. This could lead to an underestimate of P. knowlesi burden, affecting epidemiological assessments of P. knowlesi infection risk.

Our analysis is subject to several limitations. Our systematic review may not have identified all studies with relevant data, including those studies not indexed by MEDLINE and Web of Science. We limited our review to studies that tested for all five Plasmodium spp. However, studies that included subsets of the five Plasmodium spp. could still be informative of species-level diagnostic performance.

Our latent class model assumed conditional independence of diagnosis across Plasmodium spp. Future extensions of this work could account for the correlation structure of LM diagnosis across the Plasmodium spp., due to shared morphologies of the parasites. Additionally, due to the data available for this analysis, it was not possible to estimate sensitivity and specificity as a function of Plasmodium parasitemia. Future work could estimate the relationship between diagnostic performance and parasitemia, should such data be made available. Finally, to address nonidentifiability issues, we assumed PCR to be the gold standard. However, varying limits of detection and the complexity of multiplex assays due to amplification issues, parasitemia, and laboratory settings among other factors may weaken this assumption. Although sensitivity analyses revealed that our
species-level estimates were robust to this assumption, future work could jointly estimate the diagnostic performance of both methods, providing a more complete characterization of misdiagnosis of Plasmodium spp. in settings with co-circulation of multiple Plasmodium species.

This analysis estimates that the diagnostic performance of light microscopy for Plasmodium spp. may be limited in settings of cocirculation. Such misdiagnosis could have implications for malaria treatment and control, so more sensitive diagnostics, such as PCR, may be warranted in settings where Plasmodium spp. parasites routinely co-circulate.

Received November 4, 2021. Accepted for publication July 5, 2022. Published online December 12, 2022.


Acknowledgments: We thank Parker Ladwig for helpful advice on conducting a systematic review.

Financial support: This work was supported by a Graduate Research Fellowship from the National Science Foundation to J. H. H., a Richard and Peggy Notebaert Premier Fellowship from the University of Notre Dame to J. H. H., the National Institute of General Medical Sciences (grant no. 1R35GM143029-01) to T. A. P., and the National Institute of Allergy and Infectious Diseases (grant no. R21-AI137891) to J. H. H.

Disclaimer: The funding sources had no role in the study design, data collection, analysis, interpretation, or writing of the report.

Availability of data and materials: All code and data to reproduce the analyses can be accessed at https://github.com/johnhhuber/malaria_misdiagnosis.git.

Authors’ addresses: John H. Huber, Margaret Elliott, Cristian Koepfli, and T. Alex Perkins, Department of Biological Sciences, Notre Dame, IN, E-mails: huber.j@wustl.edu, melliott5@nd.edu, ckoepfli@nd.edu, and taperkins@nd.edu.

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC-BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

REFERENCES


13. Cooper DJ et al., 2020. Plasmodium knowlesi malaria in Sabah, Malaysia, 2015–2017: ongoing increase in incidence despite...


