

マラリア撲滅に向けた取り組み：血液検査データの機械学習によるマラリア スクリーニング機能の開発

Efforts to Eliminate Malaria:

Developing a Malaria Screening Capability Using Machine Learning on Blood Test Data

中谷 仁

NAKATANI Hitoshi

現在もマラリアは多くの国で流行している。2023年の感染者数は2億6,300万人、死亡者数は59万7,000人に上っており^[1]、SDGsのターゲットの一つとして、2030年までのマラリア撲滅が掲げられている^[2]。重篤化や感染拡大の防止のためにはマラリア感染の早期発見による早期治療が必要であるが、初期症状が軽度であることが多く、発見が遅れてしまうことがある。もし臨床現場にて一般的に行われる血液検査によって、マラリアのスクリーニング(病気の可能性を示唆する、確定診断ではない検査)が可能になれば、見逃しを減らすことができる。そこで、自動血球計数装置Microsemi LC-667G CRPにおいて、全血球計算(以下CBC)とC反応性蛋白(以下CRP)定量を用いて行うスクリーニング機能を開発した。本稿ではマラリアが血液に及ぼす影響と、データを利用してその傾向を検出する機械学習モデルの開発について述べる。

Malaria remains endemic in many countries. In 2023, it affected 263 million people and caused 597,000 deaths^[1]. The Sustainable Development Goals (SDGs) include malaria elimination by 2030 as one of their key targets^[2]. Early detection and treatment are essential to prevent severe disease progression and further transmission. However, because initial symptoms are often mild, diagnosis may be delayed. To address this, we have developed a screening function—intended to suggest the possibility of disease rather than provide a definitive diagnosis—using complete blood count (CBC) and C-reactive protein (CRP) measurements obtained from the Microsemi LC-667G CRP, an automated hematology analyzer. This paper describes the hematological effects of malaria and the development of a machine learning model to identify malaria-related patterns from CBC and CRP data.

はじめに

マラリアは主に熱帯地域で流行している感染症である。ハマダラカに刺されることでマラリア原虫と呼ばれる単細胞寄生虫が血液を介して肝細胞へ侵入し、破壊する。その後、赤血球へ侵入し増殖、分裂を繰り返した後に赤血球を破壊する。その際に赤血球外に放出され、新たな赤血球に対して同様のサイクルを繰り返す^[3]。人に感染してマラリアを発症するマラリア原虫は5種類あり、中でも熱帯熱マラリア原虫と三日熱マラリア原虫が大きな脅威となっている。熱帯熱マラリア原虫は、アフリカ大陸で最も蔓延し、かつ生命を脅かしているマラリア原虫であり、三日熱マラリア原虫は、アフリカ以外のサハラ以南の国で主要なマラリア原虫である^[4]。症状としては発熱・貧血・脾腫(脾臓が腫れて大きくなる状態)を伴うことが多い^[5]。初期症状は軽

度であることが多く、他の熱疾患と区別することが難しいことがある。治療せずに放置すると、24時間以内に重篤化し、死に至る場合もあるため^[4]、早期発見による早期治療が重要である。

一般的なマラリア診断としては、染色した血液標本の顕微鏡観察や、マラリア原虫由来の抗原・酵素を検出する迅速診断キットによる診断(以下RDT)が行われる^[6]。しかし、比較的軽症な患者の場合や医師の臨床判断の熟練度によっては、診察において感染が疑われず検査されないことによる見過ごしが懸念される。また、最も精度の高い血液標本の顕微鏡観察による診断には多くの時間を要するため、全ての検体を診断することもできない。その課題の解決策として、発熱時に一般的に行われる血液検査のみで、マラリアのスクリーニングを行うことが考えられる^[7]。また、診断

工数や属人性の観点から、血液サンプリング後は自動で処理されるよう、自動血球計数装置においてマラリアの自動スクリーニング機能が搭載されることが望まれる。

自動血球計数装置を用いたマラリア診断の研究としては、CBC及び白血球5分類(好中球, リンパ球, 単球, 好酸球, 好塩基球の5種類, 以下5Diff)を用いて検出することで診断する方法などが報告されている^[7]。HORIBAグループでもCBC及び5Diffを用いたマラリア判定フラグを出力する機能を持った製品を販売している。一方、CBC及び白血球3分類(リンパ球, 単球, 顆粒球の3種類, 以下3Diff)を用いた研究は少ない^[8]。マラリア感染拡大地域では測定コストを抑えることが求められるため、装置価格がリーズナブルなCBC及び3Diffを測定する製品が適している。よって、CBC及び3Diffの測定からマラリアの疑いをフラグとして表示する機能(以下マラリアスクリーニングフラグ)を開発した。

Microsemi LC-667G CRPについて

Microsemi LC-667G CRP (以下LC-667G)は臨床現場で扱える自動血球計数装置である。CBCに加えて3DiffとCRPの計測を行える。コンパクトな筐体とリーズナブルな価格を実現しており、CBCと3Diffの結果は約1分で、CRP測定の結果も合わせる場合は約4分で測定完了する。

LC-667Gでは白血球数(以下WBC)、赤血球数(以下RBC)、ヘマトクリット値(以下Hct)、血小板数(以下PLT)の算出に電気抵抗法を用いている(**Figure 1**)。電気抵抗法とは生理食塩水に浮遊させた血球をアパーチャと呼ぶ微細孔に定電流を流しながら通過させ、その時に生じる抵抗変化を電圧変化として計測する方法である。血球通過時の電圧変化は血球の容積と比例するために、その電圧パルスから血球数と血球容積が得られる。この電気抵抗法は、粒子サイズの計測において高い精度を持ち、容積も測定できることから、現在も血球計数の標準法となっている。

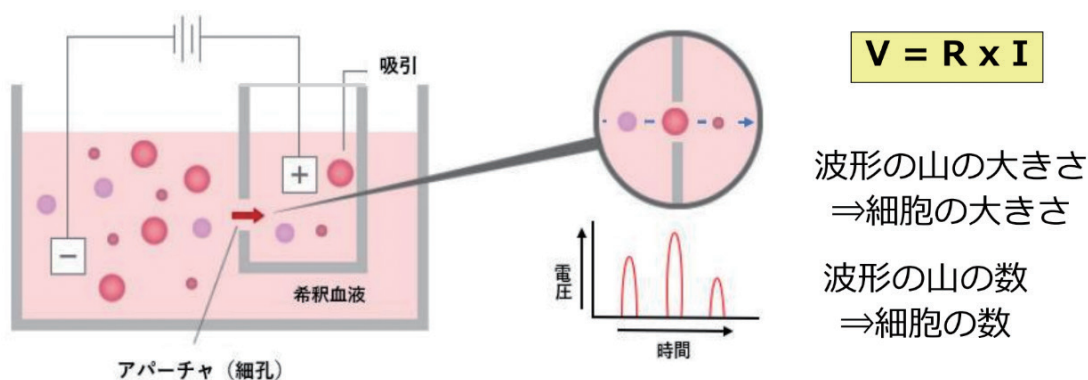
また、CRP濃度は、ラテックス免疫比濁法を用いて計測している。全血を溶血剤で溶血させた後、安定化液の存在下で抗ヒトC-反応性蛋白抗体感作ラテックス試薬を反応させると、検体中のCRPと試薬中のラテックス粒子は抗原抗体反応を起こし、ラテックス粒子が凝集する。この凝集反応による濁度変化速度を赤色光で計測し、あらかじめ標準血清を基に作成されている多項式検量線から、溶血検体中のCRP濃度を求める。同時に計測したHctから、被検体の血漿中CRP濃度に換算している。

予測因子の検討

マラリア感染の予測因子として、どの測定項目を使用すべきかを検討した。マラリアが血液に与える影響がどのように観測されるかを、データ解析と血液学的観点の両方の視点から検討した。

まず、マラリアスクリーニングフラグを開発するためのデータ収集を行った。インドのクリニックにおいて、LC-667Gを用いた血液検査を2018年から2019年のモンスーン期に実施した。発熱している全ての患者に対して血液検査を行うことにより、CBCと3DiffとCRPの測定結果を収集した。同時に、検体がマラリア陽性及び陰性であるかを判断するため、血液標本の顕微鏡観察による検査とRDTによる検査を行った。

次に、取得したデータに関してマラリア陽性検体の傾向を解析していく。まず一つに、WBCの体積度数分布に表れる異常ピークがある。これは三日熱マラリアの場合に出現しやすい(**Figure 2**)。このピークは白血球によるものではなく、主にマラリア原虫が凝集したものを検出した結果だと考えられる。一方、熱帯熱マラリアの場合では、このピークが三日熱マラリアの場合よりも小さくなる。これは大きなサイズの熱帯熱マラリア原虫は脾臓に隔離されやすく、末梢血に存在しづらいためである^[8]。



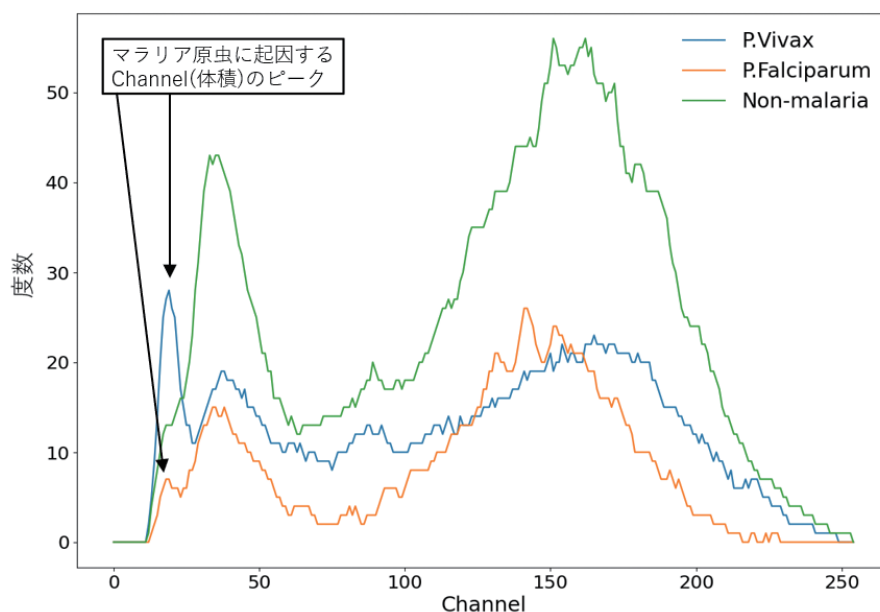


Figure 2 Volume distribution of white blood cells in samples of *P. Vivax*, *P. Falciparum* and Non-malaria. The unit of the horizontal axis, Channel, takes values from 0 to 255 that are proportional to the volume of white blood cells.

次にCRP濃度の上昇が確認できた。CRP濃度は炎症のバイオマーカーであり、感染症などで値が上昇することが知られている。マラリアに特異的なマーカーではないが、マラリアの重症度を評価するマーカーとして報告されている^[9]。また、マラリアとデング熱の鑑別にもCRP濃度は有用である。この2つは流行地域の多くが重なっているため鑑別が重要であるが、症状もCBCの測定結果も似た傾向を示すために鑑別が困難である。一方、Figure 3のように、デング熱よりもマラリアの方がCRP濃度は高い傾向にある。これは、デング熱の病原体はデングウィルスであり、ウィルス感染症ではCRP濃度の上昇は軽度であるためである^[10]。

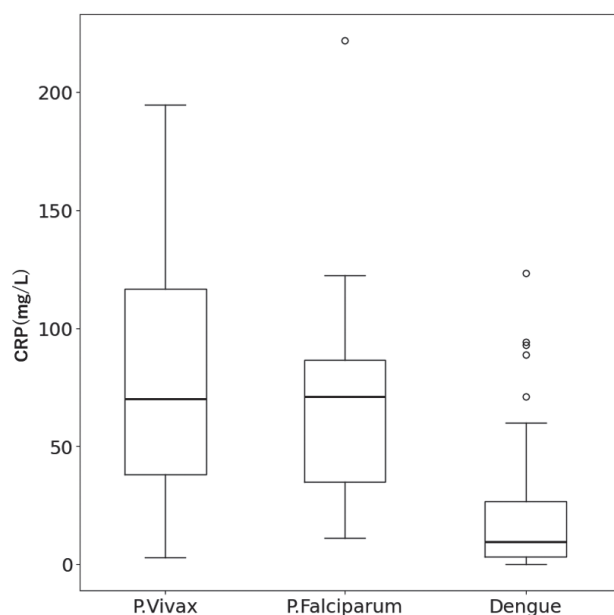


Figure 3 Comparison of CRP concentrations in samples of *P. Vivax*, *P. Falciparum* and dengue fever.

また、PLTの減少も確認された。PLTの減少については、様々なメカニズムが関与していると推察される。例えば、脾腫によって脾臓内に通常よりも多くの血小板が蓄えられてしまい末梢血の血小板が少なくなることや、マラリア原虫による赤血球の破壊に伴い、血小板が凝集し大きな一つの血球として数えられることで血小板数が少なく計測されることなどによって起こる^[11]。

CRP及びPLTが他の測定値と比較しマラリア感染に対する鑑別能が高いことを確認するために、Figure 4とTable 1を示す。Figure 4はマラリア陽性と陰性を鑑別する際のROC Curve(Receiver Operating Characteristic Curve)である。ROC Curveとは、各値に対するカットオフ値(陽性及び陰性を判断するための閾値)を変更した場合の“Sensitivity”(陽性検体の内正しく陽性と判断する割合)と“1-Specificity”(陰性検体を間違って陽性と判断する割合)とをプロットしたものである。前者は高い程良く、後者は低い程良いため、グラフが図の左上に向かっていく程良い鑑別能を持った値である。これを定量化したものがAUC(Area Under the Curve)であり、その名の通りグラフの下面積を表している。Figure 4とTable 1より、各値単体での鑑別能としてはCRP及びPLTが、WBCやMCHC(平均赤血球ヘモグロビン濃度)、RDW(赤血球粒度分布幅)よりも優れていることが分かる。

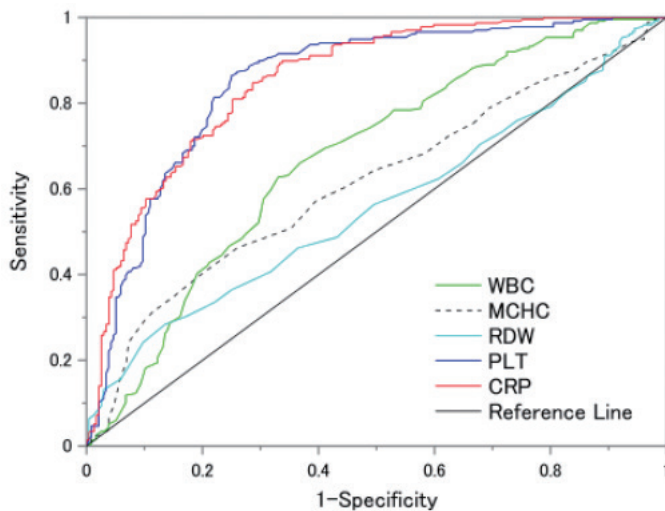


Figure 4 Comparison of ROC Curves for hematological parameters in the malaria diagnosis

Table 1 Comparison of AUCs for hematological parameters in the malaria diagnosis

	AUC
WBC	0.672
MCHC	0.613
RDW	0.558
PLT	0.849
CRP	0.852

モデリング

マラリアスクリーニングフラグを算出する手法の中核には機械学習アルゴリズムにより作成されたマラリアスクリーニングモデル(以下モデル)が用いられている(**Figure 5**)。この機械学習では、要求を満たすための2つの工夫を行っている。

まず、モデルの解釈可能性・説明可能性が担保できることである。モデルの解釈可能性とは、機械学習によって作成されたモデルが入力からどのようにして出力を導き出したかを説明できるかどうかの程度であり、モデルの説明可能性とは、出力を導出した理由を説明できる程度である。これらが十分でないものは医師が患者に診断理由を説明できないため、臨床現場で採用されないと考えた。今回、機械学習アルゴリズムとしては線形スパースモデリングを採用

しており、この機械学習アルゴリズムで作成されるモデルでは式1のように各入力とそれらに対応する重みとの積和演算により出力を計算しているため、各入力の寄与度が定量化できる解釈可能なモデルである。また、この機械学習アルゴリズムには、入力の中から重要な因子を自動で絞りこみ選択する機能があり、選択された因子のみを使用することでモデルの説明可能性を向上させている。前章で挙げているように、モデルに入力される因子も、マラリアが血液に及ぼす機序を考慮して設計されている。かつ各因子が、マラリア感染の疑い度合いと相関があるため、出力への影響を理解することも容易である。

$$Y = \sum_{i=1}^n W_i X_i \quad (1)$$

ここで、 X_i は*i*番目の入力、 W_i は X_i に対する重み、 Y は出力であるマラリア感染の疑い度合いである

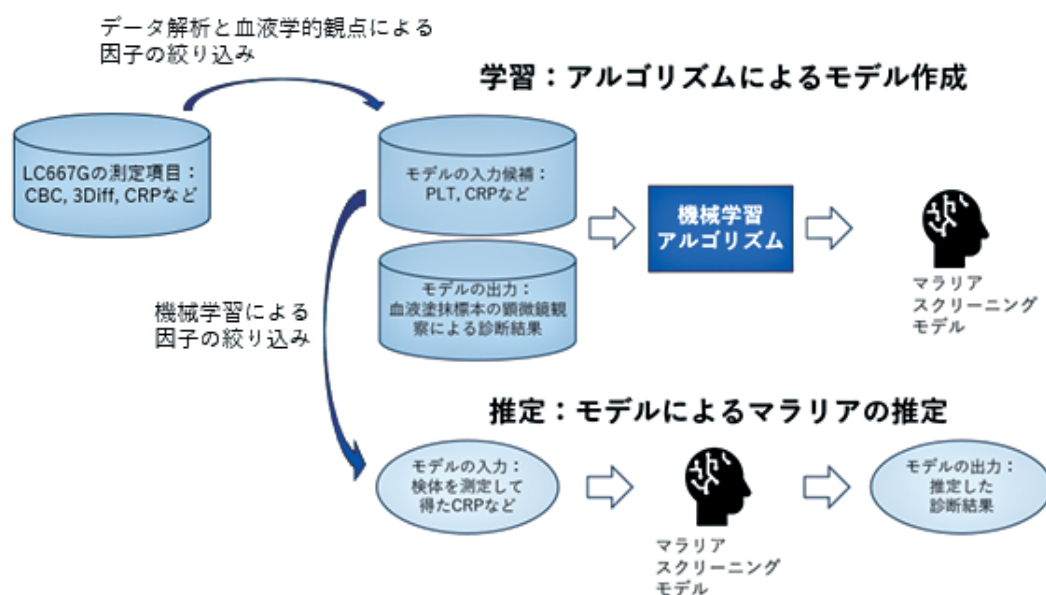


Figure 5 Workflow for the development and application of a machine learning model for malaria screening based on hematological parameters

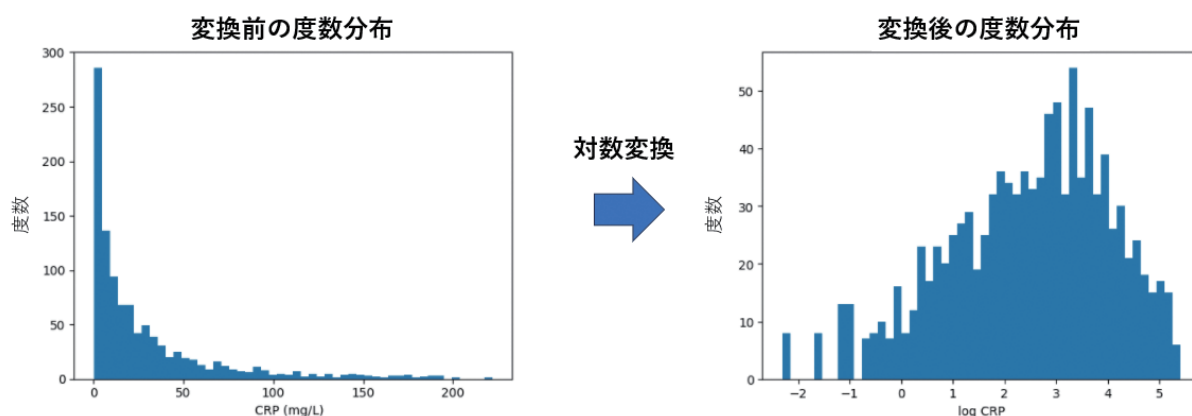


Figure 6 Comparison of CRP value distributions before and after log transformation, showing improved normalization

次に、精度が十分に高いことである。スクリーニングフラグを信頼するためには、マラリア診断のゴールドスタンダードである血液標本の顕微鏡観察での診断との相関が高くなければならないため、以下のような工夫を行っている。今回使用した機械学習アルゴリズムでは、入力された因子が正規分布であることを仮定して動作するため、実際のデータの分布が乖離しているとモデルの精度が落ちる。例えば、CRPは右歪曲分布になっており、対数変換をすることで正規分布に近づけている(**Figure 6**)。また、マラリアの発症数のピークとなるモンスーン期に採取されたデータとは言え、マラリア陽性の検体数は陰性の検体数と比較して少ない。一般的に機械学習ではデータ数が多い側の出力の傾向を強く学習してしまうため、各データの学習時の重みを変化させて学習の度合いをコントロールしている。また、スクリーニング用途ではSensitivityが高いことが要求されるため、学習方法を調整し、ROC Curveを描画し要求精度を満たす最適な判定閾値を求めるようにしている。また、血液の傾向は人種・年齢・性別などが大きく影響するため、モデルが使用される範囲を検討し誤った判断をしてしまうリスクを低減している。これらにより、モデルの精度を向上させた。

おわりに

CBCと3DiffとCRPによるマラリアスクリーニングフラグを開発した。このフラグに関するモデルには機械学習を用いており、非属人的で高精度かつ説得力のある判断が、自動でできるようになった。省力化により従来よりも多くの患者の対応が可能となり、かつ症状が顕在化していない患者に対してもスクリーニングを実施することでマラリアの早期発見に貢献できる。これからもHORIZAグループは、DX・AI技術を用いた計測・分析を通して世界中の人々の健康維持に貢献していく。

*編集局注：本内容は特段の記載がない限り、本誌発行年時点での自社調査に基づいて記載しています。

参考文献

- [1] World Health Organization, "World malaria report 2024"
<https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2024>
(2025/04/20参照)
- [2] 日本SDGs協会 "SDGsについて"
<https://japansdgs.net/about/sdgs/>
(2025/07/17参照)
- [3] Baer K, Klotz C, Kappe SHI, Schnieder T, Frevert U (2007) Release of Hepatic Plasmodium yoelii Merozoites into the Pulmonary Microvasculature. *PLoS Pathog* 3(11): e171.
<https://doi.org/10.1371/journal.ppat.0030171>
- [4] World Health Organization "Malaria"
<https://www.who.int/news-room/fact-sheets/detail/malaria>
(2025/04/20参照)
- [5] マラリアにおける腸内細菌叢と病態の関連性 久枝 一 *Journal of Japanese Biochemical Society* 88(5): 649-653 (2016) 公益社団法人日本生化学会
- [6] MSDマニュアル プロフェッショナル版 "マラリア"
<https://www.msdmanuals.com/ja-jp/professional/13-%E6%84%9F%E6%9F%93%E6%80%A7%E7%96%BE%E6%82%A3/%E8%85%B8%E7%AE%A1%E5%A4%96%E5%AF%84%E7%94%9F%E5%8E%9F%E8%99%AB/%E3%83%9E%E3%83%A9%E3%83%AA%E3%82%A2>
(2025/04/20参照)
- [7] Campuzano-Zuluaga, G., Hänscheid, T., & Grobusch, M. P. (2010). Automated haematology analysis to diagnose malaria. *Malaria journal*, 9, 346.
<https://doi.org/10.1186/1475-2875-9-346>
- [8] Nishimura, J., Dharap, P., & Raimbault, S. (2021). The utility of basic blood counts, WBC histogram and C-reactive protein in detecting malaria. *BMC infectious diseases*, 21(1), 1006.
<https://doi.org/10.1186/s12879-021-06704-5>
- [9] Mu, A. K., Bee, P. C., Lau, Y. L., & Chen, Y. (2014). Identification of protein markers in patients infected with Plasmodium knowlesi, Plasmodium falciparum and Plasmodium vivax. *International journal of molecular sciences*, 15(11), 19952–19961.
<https://doi.org/10.3390/ijms151119952>
- [10] 中西雅樹, 稲葉 亨, 藤田直久. シリーズ : CRP-基礎と臨床応用 1. CRP (C reactive protein; C反応性蛋白). 医療と検査機器・試薬 35: 725-730, 2012.
- [11] Lacerda, M. V., Mourão, M. P., Coelho, H. C., & Santos, J. B. (2011). Thrombocytopenia in malaria: who cares?. *Memorias do Instituto Oswaldo Cruz*, 106 Suppl 1, 52–63.
<https://doi.org/10.1590/s0074-02762011000900007>



中谷 仁

NAKATANI Hitoshi

株式会社堀場製作所
開発本部 プロセスエンジニアリングセンター
IoT&データアナリティクス部
サブリーダー

Assistant Section Leader,
IoT & Data Analytics Dept.,
Process Engineering Center,
R&D Division
HORIBA, Ltd.