RESEARCH ARTICLE

REVISED An assessment of interobserver agreement on lesion size, morphology and clinical phenotype in cutaneous leishmaniasis caused by *Leishmania aethiopica* in Ethiopia

[version 2; peer review: 2 approved]

Amel Beshir Mohammed[1,2], Fewzia Shikur Mohammed[1,2], Feleke Tilahun Zewdu[3,4], Shimelis Doni Nigusse[1,2], Saba Lambert[1,5], Michael Marks[5-7], Stephen L. Walker[5,7], Endalamaw Gadisa[8], SHARP Collaboration

[1]ALERT Hospital, Addis Ababa, Ethiopia
[2]Department of Dermatovenereology, College of Health Sciences, Addis Ababa University, Addis Ababa, Addis Ababa, Ethiopia
[3]Boru Meda Hospital, Boru Meda, Ethiopia
[4]Department of Epidemiology and Biostatistics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia
[5]Department of Clinical Research, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK
[6]Division of Infection and Immunity, University College London, London, UK
[7]Hospital for Tropical Diseases, University College London Hospitals NHS Foundation Trust, London, UK
[8]Department of Malaria and Neglected Tropical Disease Research, Armauer Hansen Research Institute, Addis Ababa, Ethiopia

v2 **First published:** 18 Feb 2025, **5**:12
https://doi.org/10.3310/nihropenres.13869.1
Latest published: 14 May 2025, **5**:12
https://doi.org/10.3310/nihropenres.13869.2

Abstract

Introduction

Cutaneous leishmaniasis (CL) remains a major public health challenge, especially in endemic regions like Ethiopia, where an estimated 40,000 new cases occur annually. Effective treatment evaluation for CL relies on consistent clinical assessments, yet variability in lesion descriptions can complicate reliable outcome measures.

Methods

We conducted an inter-reliability study of clinicians' evaluations of CL lesion morphology and size at ALERT Hospital, Addis Ababa. Twelve clinicians independently examined 12 patients with parasitologically confirmed CL, each clinician assessing lesion morphology, size, and severity.

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 14 May 2025	✓ view	✓ view
version 1 18 Feb 2025	? view	✗ view

1. **Lucinda Clare Fuller**, International Foundatioon for Dermatology, London, UK
annika Wilder-Smith, University of Zurich, Zurich, Switzerland

2. **Maria Adelaida Gómez**, Centro Internacional de Entrenamiento e Investigaciones Medicas, Cali, Colombia

Results

We found high consistency in reporting major morphological categories (e.g., plaques) but significant variability in secondary features like dyspigmentation and scale, as well as mucosal involvement. Lesion size measurements showed limited variability, suggesting its reliability as a potential measure for future clinical trials. Disparities in severity assessments highlight the need for a standardized scoring system in CL.

Discussion

Our findings underscore the importance of training for consistent, high-quality clinical evaluations of CL and suggests that lesion size could be a reproducible outcome measure in treatment efficacy trials.

Plain Language Summary

Cutaneous leishmaniasis is a significant public health issue in Ethiopia, with around 40,000 new cases each year. The disease can cause a wide variety of different types of skin lesions which can complicate diagnosis and assessment. Conducting reliable treatment evaluation requires consistent assessment of cutaneous leishmaniasis lesions, especially in clinical trials.

We conducted a prospective study at ALERT Hospital in Addis Ababa to assess how consistently 12 clinicians, who were all experienced in managing patients with cutaneous leishmaniasis, evaluated the appearance, size and severity of cutaneous leishmaniasis lesions in 12 patients.

We found that whilst clinicians were broadly consistent in identifying the major features such as the type of lesions they showed more significant variability in describing secondary features and severity. Lesion size measurements were more consistent, making them a promising tool for future clinical trials. The findings highlight the need for better training and standardized scoring systems to improve the reliability of clinical evaluations in cutaneous leishmaniasis.

Keywords

cutaneous leishmaniasis

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Michael Marks (Michael.Marks@lshtm.ac.uk)

Author roles: Mohammed AB: Methodology, Writing – Review & Editing; **Mohammed FS**: Investigation, Methodology, Writing – Review & Editing; **Zewdu FT**: Investigation, Methodology, Writing – Review & Editing; **Nigusse SD**: Investigation, Methodology, Writing – Review & Editing; **Lambert S**: Conceptualization, Funding Acquisition, Investigation, Methodology, Writing – Review & Editing; **Marks M**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Walker SL**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Writing – Original Draft Preparation; **Gadisa E**: Conceptualization, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Review & Editing;

Competing interests: No competing interests were disclosed.

Grant information: This project is funded by the National Institute for Health and Care Research (NIHR) under its ['Research and Innovation for Global Health Transformation programme (Grant Reference Number NIHR200125)]. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Mohammed AB, Mohammed FS, Zewdu FT *et al.* **An assessment of interobserver agreement on lesion size, morphology and clinical phenotype in cutaneous leishmaniasis caused by *Leishmania aethiopica* in Ethiopia [version 2; peer review: 2 approved]** NIHR Open Research 2025, **5**:12 https://doi.org/10.3310/nihropenres.13869.2

First published: 18 Feb 2025, **5**:12 https://doi.org/10.3310/nihropenres.13869.1

REVISED **Amendments from Version 1**

Minor changes in response to peer reviewers. New Figure 1 added.

Any further responses from the reviewers can be found at the end of the article

Introduction

Cutaneous leishmaniasis (CL) is the most common form of leishmaniasis. The disease causes skin lesions on exposed parts of the body most commonly at the site of an infected sandfly bite. With or without treatment CL lesions can lead to life-long scars and result in serious disability or stigma. It is estimated that 600 000 to 1 million new cases occur worldwide annually with 40,000 new cases of CL in Ethiopia[1]. Three *Leishmania* species have been reported in Ethiopia, but the vast majority of cases are thought to be due to *L aethiopica*.

L. aethiopica is associated with a variety of clinical phenotypes. The most frequent presentation is localized CL (LCL), but infection is also associated with mucocutaneous leishmaniasis (MCL), and diffuse CL (DCL). LCL usually starts as a papule that gradually enlarges forming nodules or plaques. Ulceration may occur with *L. aethiopica*. MCL can occur either via the bite of the sand fly at the border of mucosal surfaces or through extension from an adjacent cutaneous lesion. It most commonly involves the nasal mucosa but the lips and oral mucosa can also be involved. DCL is a chronic and progressive disorder which starts with papular or nodular lesions which become larger and more numerous[2].

Despite the large number of affected individuals, studies of the clinical manifestations of *L. aethiopica*, including the morphology of the lesions, progression of the disease and response to therapy, are limited. A study in North-west Ethiopia reported the most common features to be induration, erythema, ulceration, and crusting.

The WHO strategic framework for integrated control and management of "skin-related" NTDs highlighted anti-microbial therapy for CL as a research gap and this is particularly the case for treatment of *L. aethiopica* infection[3,4]. Robust outcome measures are required for treatment efficacy studies. Outcome measures in CL in studies are based on clinical evaluation of individual lesions including re-epithelization in ulcerated lesions and flattening in non-ulcerated ones and lesion measurement. In 2018, a proposal for harmonized clinical trial methodologies for cutaneous leishmaniasis was published to develop a consensus and standardize clinical evaluation as a measure of disease response and cure in CL[5].

Given the reliance on clinical observations of lesion morphology and severity it is therefore essential to evaluate whether the clinical evaluation of individual physicians is comparable. We therefore conducted a study to assess the inter-rater reliability of clinicians' assessment of lesion morphology and size in individuals affected by CL in Ethiopia.

Methods

Patient and Public Involvement

Neither patients nor the public were directly involved in the design or analysis of the study.

The study was conducted nested within a larger cohort study of CL in Ethiopia. As part of this larger study, we conducted an inter-observer exercise at ALERT comprehensive specialized Hospital, Addis Ababa[6].

Examiners were selected from the clinical staff of ALERT Hospital and members of the Skin Health Africa Research Programme (SHARP) team. We collected information on their previous training in skin disease and experience of managing CL.

Clinicians at ALERT Hospital identified 12 individuals with a parasitologically confirmed diagnosis of CL who were invited to participate in the study. A standardized proforma was used to capture routine clinical data about each affected individual. For individuals with more than one clinical lesion of CL, the treating clinician indicated the "index" lesion to be assessed in the inter-observer assessment exercise. This "index" lesion was the lesion identified by the affected individual as the one of greatest concern prior to the start of the exercise.

A computer generated latin-square was used to randomize an order of evaluation for each observer and affected individual. Each observer was assigned a unique letter and each affected individual a unique number. To measure inter-observer agreement each observer independently examined all 12 individuals with CL once for a period of 5 minutes.

Examiners recorded on a standard data collection sheet their assessment of each "index" lesion. Standardized definitions of CL developed for our cohort study (Box 1) were used[6]. For each lesion examiners were asked to record morphology, size and the presence or absence of clinical features including induration, ulceration, crust, hyper/hypopigmentation, scarring and mucosal involvement. Examiners measured the largest diameter of the lesion using a disposable tape measure. Examiners were asked to provide an assessment of the clinical phenotype of CL and to classify the disease severity as mild or moderate or severe. All data were double entered into an electronic study database for analysis.

Box 1. Operational definitions of Cutaneous Leishmaniasis (CL).

- CL is diagnosed in a person with skin and/or mucosal lesions with evidence of *Leishmania* infection in the affected tissues characterised by the presence of *Leishmania* amastigotes on smear microscopy or growth of *Leishmania* promastigotes in culture or the detection of *Leishmania* DNA by polymerase chain reaction.
- Localized CL: A confirmed case of leishmaniasis, with no mucosal involvement, characterized by ten or fewer cutaneous papules and/or nodules and/or plaques with or without ulceration involving one body site

- Multi-regional localized CL: A confirmed case of leishmaniasis, with no mucosal involvement, characterized by ten or fewer cutaneous papules and/or nodules and/or plaques with or without ulceration involving two or more body sites
- Mucocutaneous leishmaniasis: A confirmed case of leishmaniasis characterized by ten or fewer papules and/or nodules and/or plaques with or without ulceration involving skin and an adjacent mucosal surface
- Pure Mucosal leishmaniasis: A confirmed case of leishmaniasis characterized by papules and/or nodules and/or plaques with or without ulceration involving exclusively a mucosal surface
- Diffuse CL: A confirmed case of leishmaniasis characterized by eleven or more papules and/or nodules and/or plaques with or without mucosal involvement.

Body sites are classified as:

1. Head and neck
2. Torso - anterior (including genitalia)
3. Torso - posterior (including buttocks)
4. Right upper limb
5. Left upper limb
6. Right lower limb
7. Left lower limb

As there is no published data on the degree of agreement in assessing cases of Cutaneous Leishmaniasis we used a convenience sample size based on ensuring we had a range of both examiners and patients with clinical differing clinical phenotypes. The primary planned analysis was descriptive. We report the proportion of affected individuals judged to have each feature by the examiners. For lesion size we describe variability by reporting the median and IQR lesion size reported for each index lesion. Findings are reported in line with the STROBE guidelines. Analysis was conducted in R version 4.2.1.

Results

Twelve affected individuals agreed to participate in the study. Seven were male and five female. The median age was 25 years. Based on the treating clinician's assessment six had been diagnosed with LCL, four with MCL and two with pure mucosal CL. All twelve individuals were currently receiving inpatient treatment with intramuscular sodium stibogluconate. Clinical photos of CL lesions are shown in Figure 1.

Twelve examiners took part in the study. Seven of the clinicians were male and 5 were female. The median age of examiners was 39. The majority (n = 10, 83%) were clinicians with specialist postgraduate training in dermatology and six (50%) were accredited dermatologists. (Table 1).

Some clinical features appeared to be more consistently reported by examiners than others (Table 2). For lesion morphology all examiners agreed on the presence of plaques in one individual, whilst 10 or more examiners agreed on the presence of plaques in five. In the remaining individuals with CL either eight or nine examiners assessed the lesion to be a plaque. All examiners agreed there were no nodular lesions in eight individuals. For the remaining four affected individuals one to four examiners felt there were nodular lesions. There was greater variability in reporting of the presence or absence of papules. For five individuals all examiners agreed papules were absent whereas for seven individuals between one to five examiners assessed they were present.

Figure 1. Clinical Photos of patients assessed in the study.

Table 1. Characteristics of individuals conducting the examinations.

Median Age (years)	39
Male	7
Years experience	8.3 (SD 6.8)
Clinical specialty	
Dermatologist	6
Dermatology Trainee	3
Infectious Diseases	1
MSc in Dermatology	1
GP with Dermatology postgraduate diploma	1

Involvement of the mucosa was reported by all examiners for three individuals and by 11 examiners in two others. For three individuals with CL all examiners agreed there was no mucosal involvement and for a further two individuals only one or two examiners assessed the mucosa to be affected. For two individuals there was more marked disagreement with seven and nine examiners reporting involvement and five and three examiners reporting no mucosal involvement.

For lesion features all examiners agreed on the presence of erythema in three individuals and at least 10 examiners agreed it was present in a further four. For the remaining individuals with CL there was marked variability with between three to eight examiners reporting erythema. For five individuals all examiners reported the presence of induration and for a further four at least 11 examiners felt induration was present. For

Table 2. **The number of times each clinical feature was reported for each patient.**

Affected Individual ID	Treating Clinician Phenotype	Erythema	Indurated	Ulcerated	Crusted	Scaly	Scar	Hypopigmented	Hyperpigmented	Papule	Nodule	Plaque	Mucosa
1	LCL	3	12	2	11	5	4	1	5	5	0	11	0
2	LCL	7	11	0	12	11	0	1	10	0	2	8	7
3	Pure Mucosal	11	12	1	1	10	0	1	0	1	1	8	12
4	MCL	12	11	9	2	6	1	0	2	1	0	9	11
5	MCL	12	12	0	8	10	1	5	5	0	0	10	11
6	MCL	12	12	4	7	1	6	5	8	2	0	11	12
7	LCL	6	12	0	0	2	11	0	10	5	0	11	2
8	LCL	5	11	0	2	10	2	1	8	0	0	12	0
9	LCL	10	9	0	3	11	0	1	4	2	0	10	1
10	Pure Mucosal	11	9	1	0	9	1	3	1	0	0	9	12
11	LCL	8	9	2	12	2	0	0	5	0	3	9	0
12	MCL	11	11	0	10	10	2	0	8	3	1	9	9

NIHR Open Research 2025, 5:12 Last updated: 03 JUL 2025

the remaining three affected individuals nine examiners felt induration was present and three felt it was absent.

For six individuals with CL all examiners agreed that there was no evidence of ulceration. For the remaining six patients the number of examiners reporting ulceration varied from one to nine. For two individuals all 12 examiners agreed that there was evidence of crust and for two individuals all examiners agreed crust was absent. For the remaining eight individuals with CL the number of examiners reporting crust varied from one to eleven.

There were no individuals with CL whom examiners were in agreement on the presence or absence of scale but there were six where at least 10 examiners agreed scale was present and three where at least 10 examiners agreed it was absent.

Finally, all examiners agreed there was no evidence of scarring in four individuals. For the remaining eight between one and eleven examiners deemed scarring to be present.

Estimates of the reported maximum lesion diameter were broadly consistent across examiners but graphical visualization with a modified Bland-Altman plot suggests variation increased with the size of the lesion (Figure 2)[7].

The majority of observers assessed that two third of patients had moderate to severe disease with the remaining affected individuals assessed as having mild-moderate disease. There were only two individuals with CL in whom there was a 2-grade discrepancy in the assessment of severity but there were no individuals in whom all examiners assigned the same severity score and only one in whom there was more than 80% agreement. Disagreement about mucosal involvement also meant that there were only four (33%) individuals with CL where all 12 examiners and the treating clinicians agreed on the clinical phenotype (Table 3).

Discussion

In the first ever study evaluating the inter-observer reliability of clinical examination to assess lesion morphology in CL we demonstrated considerable variation in the reporting of several key clinical features. Reporting of overall lesion morphology as plaque, nodule or papule was generally highly consistent but the presence or absence of secondary clinical characteristics such as dyspigmentation or the presence of scale were more variable. Lesion size is proposed as a key characteristic for assessing treatment outcomes in future trials[5]. Reassuringly, we showed that there was limited variation in the assessed size of lesions when measured by multiple independent assessors, although there seemed to be evidence that variability was highest for larger CL lesions. Overall, these data suggest that lesion size assessment is likely to be a reliable measure in both observational and interventional prospective studies.

We found assessment of mucosal involvement to be variable. As this has important implications for treatment and characterization for clinical trials it may be advisable in some cases to ensure clinicians are well trained in assessing for mucosal involvement, and in some circumstances obtaining an assessment by an ear, nose and throat specialist to allow accurate characterization. There is not yet a standardized definition of severity for CL[6]. In the absence of such a system we found little consensus amongst clinicians when assessing individuals

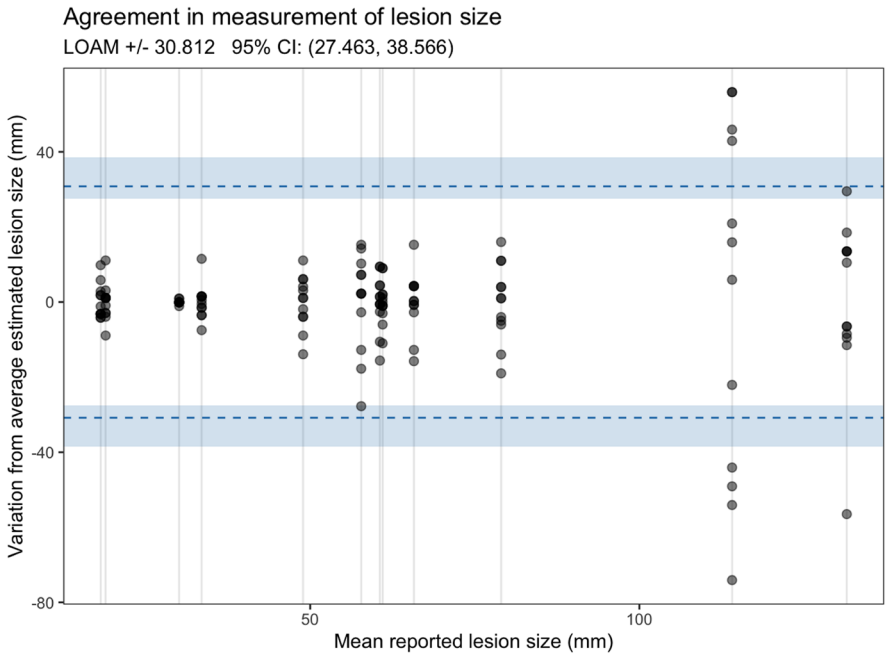

Figure 2. Bland-Altman plot showing variation in assessment of lesion size.

Table 3. Clinical Phenotype and Severity as judged by the examiners.

Patient	Treating Clinician Phenotype	Phenotype Assessment				Severity Assessment		
		DCL	LCL	MCL	Pure Mucosal	Mild	Moderate	Severe
1	LCL	0	12	0	0	0	8	4
2	LCL	0	5	7	0	1	7	4
3	Pure Mucosal	0	0	5	7	0	5	7
4	MCL	0	2	10	0	0	7	5
5	MCL	0	2	10	0	0	2	10
6	MCL	0	0	12	0	0	3	9
7	LCL	1	9	2	0	0	5	7
8	LCL	0	12	0	0	7	4	1
9	LCL	0	11	1	0	5	6	1
10	Pure Mucosal	0	0	10	2	6	3	3
11	LCL	0	12	0	0	9	3	0
12	MCL	0	3	9	0	1	4	7

with CL which poses challenges in reliable reporting of their characteristics in both observational and interventional studies. Development of a validated severity score should be considered to help standardize practice.

Inter-observer variability is well recognized in clinical assessment particularly when assessing subjective features of disease. Similar to our findings, previous studies in other NTDs have shown that the degree of reliability may vary for different clinical features within a given disease. For example, diagnosis of scabies by healthcare workers has been shown to vary markedly depending on the severity of the infestation[8,9]. Similarly for leprosy studies have shown marked differences in inter-observer reliability in determining the presence or absence of nerve thickening or in the reliability of different methods of assessing sensory loss within skin lesions[10].

The new WHO roadmap for NTDs focuses on strengthening responses to skin-NTDs[3]. Achieving this will require considerable investment in training of healthcare workers in the diagnosis and management of a range of skin diseases including CL. Given the degree of variability between expert examiners in this study our data suggest such efforts will require considerable effort to ensure high-quality, reliable and reproducible clinical diagnosis and assessment of clinical healing parameters.

Our study has limitations. Firstly, all affected individuals were inpatients receiving care for CL at a referral centre. This is likely to have resulted in selection of individuals with more severe disease than might be seen elsewhere. Similarly, all individuals were receiving treatment and therefore their lesions might differ from untreated ones. Whilst this might affect lesion morphology, we believe it is unlikely to have affected our assessment of inter-examiner reliability. In addition, we assessed inter-observer reliability at only a single time point for each patient. Future studies evaluating inter-observer agreement at different stages of patients treatment would be particularly valuable. Whilst all clinicians had experience of caring for patients with CL we did not have data on what specific training individuals had received and how this may have impacted their assessment of lesion morphology. Developing standardized training for CL assessment would be a valuable step in improving the variability we observed in this study. The generalizability of our findings to regions of the world where other *Leishmania* species are endemic is unclear, particularly in settings where ulcerated lesions may be more common.

Assessment of inter-examiner reliability for multiple examiners is challenging. Conventionally kappa-scores are calculated to compare the performance of an examiner against a reference-standard. Multi-examiner weighted kappa scores have been developed to extend this approach to situations broadly analogous to our study design. However, both conventional and multi-examiner kappa scores have weaknesses[11]. Disagreements limited to a small number of individuals being examined can have an outsize impact on kappa scores even when there is near universal consensus on the majority individuals. We therefore opted to only present descriptive statistics. Finally, our study focuses on the assessment of individuals with known CL by experts. Further studies are needed to assess the performance of other cadres of healthcare workers in the initial evaluation of individuals with suspected CL.

Trials of novel and improved therapeutics for CL are urgently needed. To provide comprehensive data on efficacy such trials will need to combine reliable and reproducible parasitological, clinical and patient reported outcome measures. Measurement of lesion size is likely to be a critical component of such measures, and our data are reassuring that this can be reliably measured by multiple independent assessors. More work is needed to

identify and improve assessment of other clinical features in order to develop robust trial outcome measures for CL.

Ethics and consent

Ethical approval was obtained from AHRI/ALERT Ethics Review Committee (Ref: PO/23/21), the National Research Ethics Review Committee (Ref: 7/2-506/m259/35), and the London School of Hygiene and Tropical Medicine (Ref: 26421) in October 2021. Written individual informed consent was obtained from all participants (examiners and affected individuals) in the study. The study complied with the Declaration of Helsinki.

Data availability statement

Underlying data

LSHTM Data Compass: An assessment of interobserver agreement on lesion size, morphology and clinical phenotype in cutaneous leishmaniasis caused by Leishmania aethiopica in Ethiopia.
https://doi.org/10.17037/DATA.00004491[12]

The project contains the following underlying data:

- Cutaneous Leishmaniasis interobserver dataset.

Extended data

LSHTM Data Compass - An assessment of interobserver agreement on lesion size, morphology and clinical phenotype in cutaneous leishmaniasis caused by Leishmania aethiopica in Ethiopia.
https://doi.org/10.17037/DATA.00004491[12]

The project contains the following supplementary data:

- CL_Interobserver_CRF – An ODK data collection form as used in the study
- CL_interobserver Codebook – A codebook describing the data coding
- Inter-observer adult consent – a Copy of the patient information sheet and consent form

Acknowledgements

We wish to thank the individuals and communities for their participation in the work of the Skin Health Africa Research Programme.

References

1. Alvar J, Vélez ID, Bern C, *et al.*: **Leishmaniasis worldwide and global estimates of its incidence.** *PLoS One.* 2012; **7**(5): 35671.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

2. Bryceson AD: **Diffuse Cutaneous Leishmaniasis in Ethiopia. I. The clinical and histological features of the disease.** *Trans R Soc Trop Med Hyg.* 1969; **63**(6): 708–37.
PubMed Abstract | **Publisher Full Text**

3. **Ending the neglect to attain the sustainable development goals: a strategic framework for integrated control and management of skin-related neglected tropical diseases**. (accessed June 25, 2024).
Reference Source

4. van Griensven J, Gadisa E, Aseffa A, *et al.*: **Treatment of Cutaneous Leishmaniasis caused by *Leishmania aethiopica*: a systematic review.** *PLoS Negl Trop Dis.* 2016; **10**(3): e0004495.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Olliaro P, Grogl M, Boni M, *et al.*: **Harmonized clinical trial methodologies for localized Cutaneous Leishmaniasis and potential for extensive network with capacities for clinical evaluation.** *PLoS Negl Trop Dis.* 2018; **12**(1): e0006141.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Mohammed AB, Mohammed FS, Zewdu FT, *et al.*: **Protocol for a prospective observational cohort study of Cutaneous Leishmaniasis in Ethiopia [version 1; peer review: 2 approved with reservations].** *NIHR Open Res.* 2023; **3**: 49.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Christensen HS, Borgbjerg J, Børty L, *et al.*: **On Jones *et al.*'s method for extending Bland-Altman plots to limits of agreement with the mean for multiple observers.** *BMC Med Res Methodol.* 2020; **20**(1): 304.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Walker SL, Collinson S, Timothy J, *et al.*: **A community-based validation of the International Alliance for the Control of Scabies Consensus Criteria by expert and non-expert examiners in Liberia.** *PLoS Negl Trop Dis.* 2020; **14**(10): e0008717.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. Osti MH, Sokana O, Lake S, *et al.*: **The body distribution of scabies skin lesions.** *JEADV Clinical Practice.* 2022; **1**(2): 111–121.
Publisher Full Text

10. Ponnighaus JM, Fine PE, Bliss L: **Certainty levels in the diagnosis of leprosy.** *Int J Lepr Other Mycobact Dis.* 1987; **55**(3): 454–62.
PubMed Abstract

11. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics.* 1977; **33**(1): 159–74.
PubMed Abstract | **Publisher Full Text**

12. Marks M: **Data for: "An assessment of interobserver agreement on lesion size, morphology and clinical phenotype in Cutaneous Leishmaniasis caused by *Leishmania aethiopica* in Ethiopia"**. [Data Collection]. London School of Hygiene & Tropical Medicine, London, United Kingdom, 2025.
http://www.doi.org/10.17037/DATA.00004491

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 03 June 2025

https://doi.org/10.3310/nihropenres.15196.r35636

✔ **Lucinda Clare Fuller**
International Foundatioon for Dermatology, London, UK

The authors have addressed the aspects raised in the first cycle of reviews.
The manuscript is clear and concise and represents an important contribution to progressing the process of standardisation of assessment of leishmaniasis which will be vital in future trials of treatment.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a dermatologist with research experience in skin neglected tropical diseases.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 02 June 2025

https://doi.org/10.3310/nihropenres.15196.r35637

✔ **Maria Adelaida Gómez**
Centro Internacional de Entrenamiento e Investigaciones Medicas, Cali, Colombia

I have read the revised version and I think it covers the comments I made, and thus consider the manuscript acceptable for publication. The only one thing that came to mind is to ask if the authors received consent from the patients to include their full face picture in the publication? If

not, they need to have just the lesion, not the full face, so that patients can not be identified.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 02 Jun 2025

Michael Marks

All patients provided written consent for the use of their images.

Competing Interests: No competing interests were disclosed.

Version 1

Reviewer Report 12 April 2025

https://doi.org/10.3310/nihropenres.15071.r35115

Maria Adelaida Gómez
Centro Internacional de Entrenamiento e Investigaciones Medicas, Cali, Colombia

Major comments:

1. It is not clear at what time over the course of disease and treatment were the patients evaluated. In the discussion section one could infer patients were undergoing treatment, but it was not stated at what moment over the course of treatment were the patients evaluated. This leads me to question the relationship between the need (assessment of clinical cure) and the evaluation time point. Why was the study conducted on a single time point, instead of evaluating the disease progression over the course of treatment, for example?
2. For a study like this, images of lesions illustrating the characteristics that were more consistently identified vs. those that were not are necessary. This could also serve as a tool for reference to other clinicians experiencing similar difficulties in clinical assessment of CL (especially non-ulcerated lesions).
3. In the discussion the authors mention “There is not yet a standardized definition of severity for CL”. WHO and PAHO have generated guidelines that allow characterization of complicated vs. uncomplicated CL, based on number of lesions, lesion size, anatomical site of lesions, among others. This classification is the basis for selection of local vs. systemic treatments. This can be used as a baseline for definition of severity.

Minor comments

1. I suggest caution with the statement of lesions developing at the site of sandfly bite, as this is not always the case ("The disease causes skin lesions on exposed parts of the body at the site of an infected sandfly bite")
2. First paragraph of results section "mucosal CL". In the box the classifications are either mucocutaneous or mucosal leishmaniasis. In this case, what are the authors referring to? MCL or CL? Also table 2 has "Pure mucosal", please keep consistency in the use of terms.
3. What instrument was used for measurement of lesion size? This is relevant as the instrument used may introduce bias in the variability, or lack of, in this measurement among observers.

References

1. Mohammed A, Mohammed F, Zewdu F, Nigusse S, et al.: An assessment of interobserver agreement on lesion size, morphology and clinical phenotype in cutaneous leishmaniasis caused by Leishmania aethiopica in Ethiopia. *NIHR Open Research*. 2025; **5**. Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

No

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Human Cutaneous Leishmaniasis

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 26 February 2025

https://doi.org/10.3310/nihropenres.15071.r34785

?

Lucinda Clare Fuller
International Foundatioon for Dermatology, London, UK
annika Wilder-Smith
Faculty of Medicine, University of Zurich, Zurich, Switzerland

This manuscript reports a study nested within a larger piece of work that is referenced. It describes the variability in clinical findings between 12 clinicians in assessing 12 known cutaneous leishmaniasis cases. The setting is a specialist hospital in Ethiopia with acknowledged experience in managing leishmaniasis and hosting the country's flagship dermatology training program. The overarching aim of the report is to highlight the need for
1. standardised reporting structure for cutaneous leishmaniasis to support future studies developing or assessing new therapeutic approaches.
2. Development of a standardised severity score and
3. Development of standardised training for clinical assessors involved in future studies to ensure consistency.
Whilst the training and exposure of the majority of the clinicians will be "relevant" it is not standardised and it might be worth emphasising in the discussion this fact and including a recommendation for ensuring any clinical assessors in cutaneous leishmaniasis studies have undertaken specific training on lesion analysis. It might be useful to state if any such training exists (eg on the Open WHO platform NTD channel - obviously under migration to WHO academy at the moment), is this fit for purpose ie generalisable across the globe, or is there an educational gap?
I assume there is no scope for inclusion of clinical photographs to illustrate the cases and phenotype characteristics.
Specifically small comments:
Introduction: para 3 you mention a study in North west Ethiopia reporting common features - can you reference this?

Methods:
you mention a standardised proforma - is it worth showing this?
Results

Is it worth stating that although the clinicians included were experienced you are not able to specifically comment on cutaneous leishmaniasis training or else report on how this is covered in the local Ethiopian Dermatology Curriculum?

Discussion
Whilst noting the phenotypes and clinical features being measured, might you consider highlighting those most likely to be of clinical relevance in measuring treatment impact?
Mucosal involvement: whilst you might want to involve ENT in assisting the development of training tools to guide clinical assessors as to whether mucosal involvement is present or absent, it should be fairly straightforward to clarify with appropriate training.

The comments are minor and I hope helpful.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: I am a dermatologist with research experience in skin neglected tropical diseases.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.