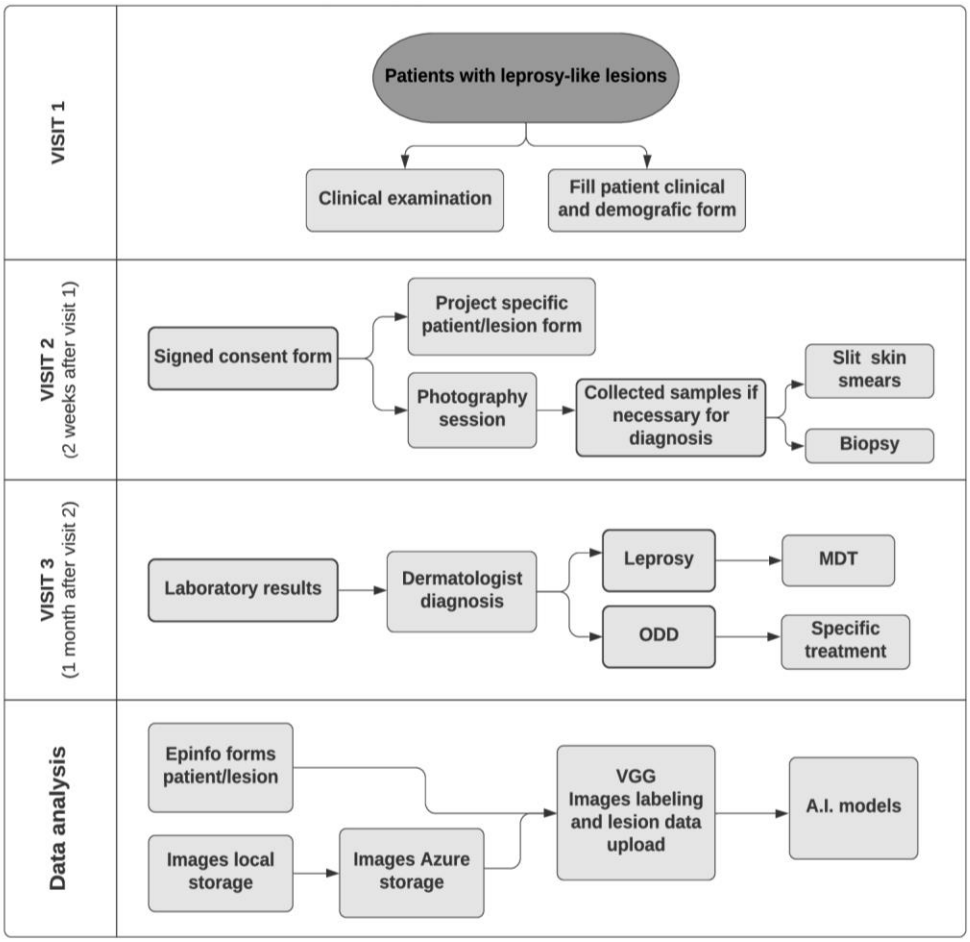
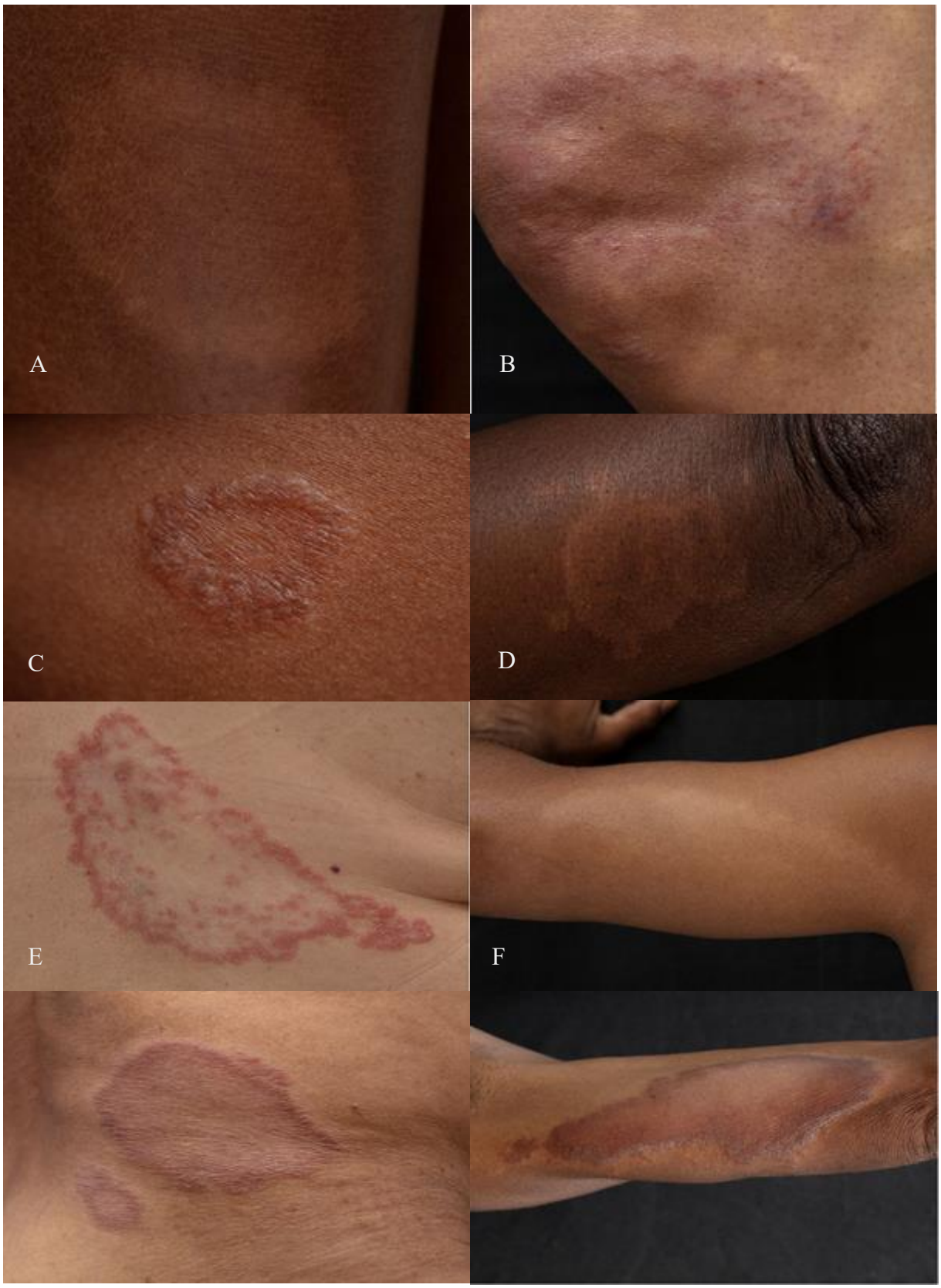
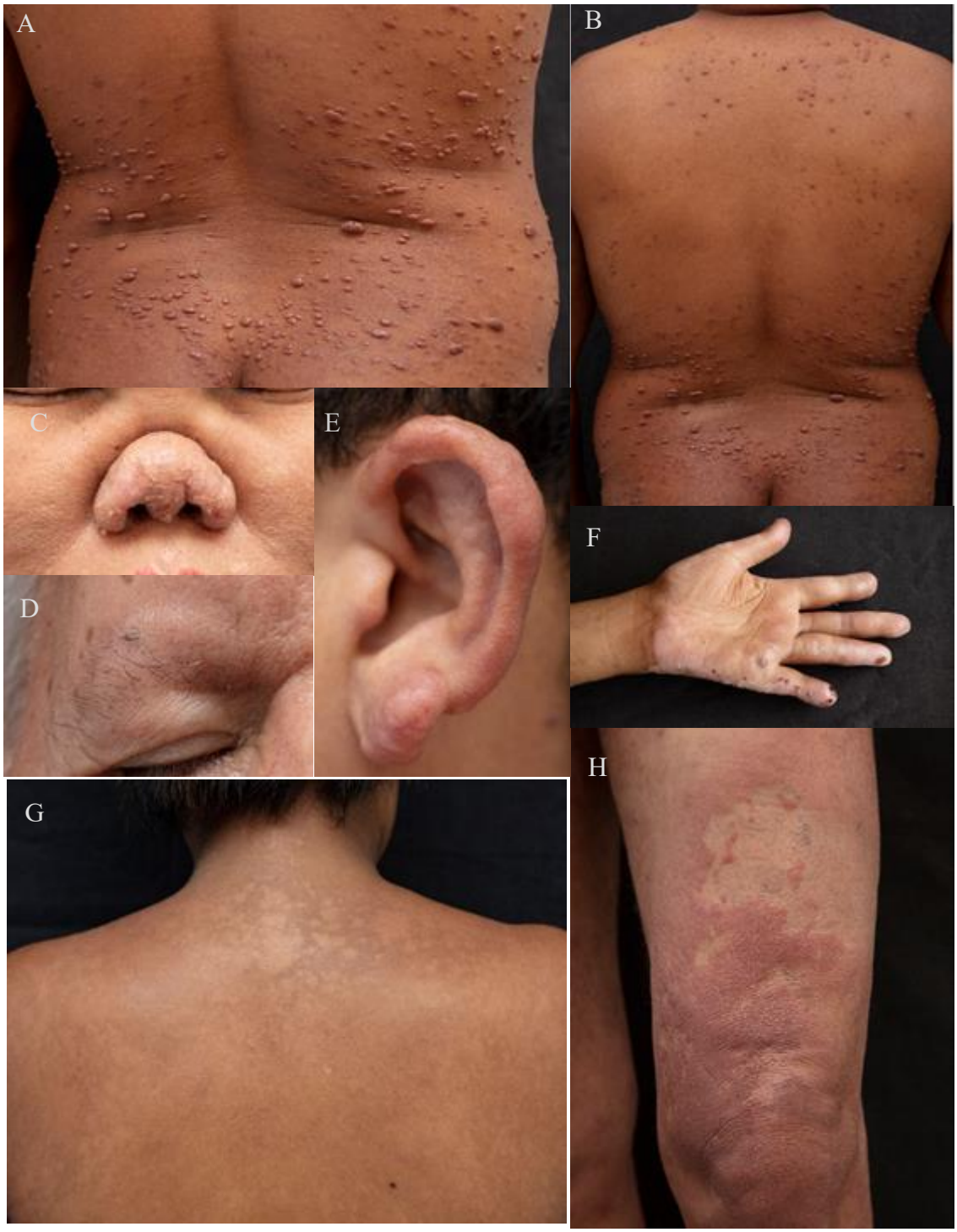
Supplementary Table of Contents:



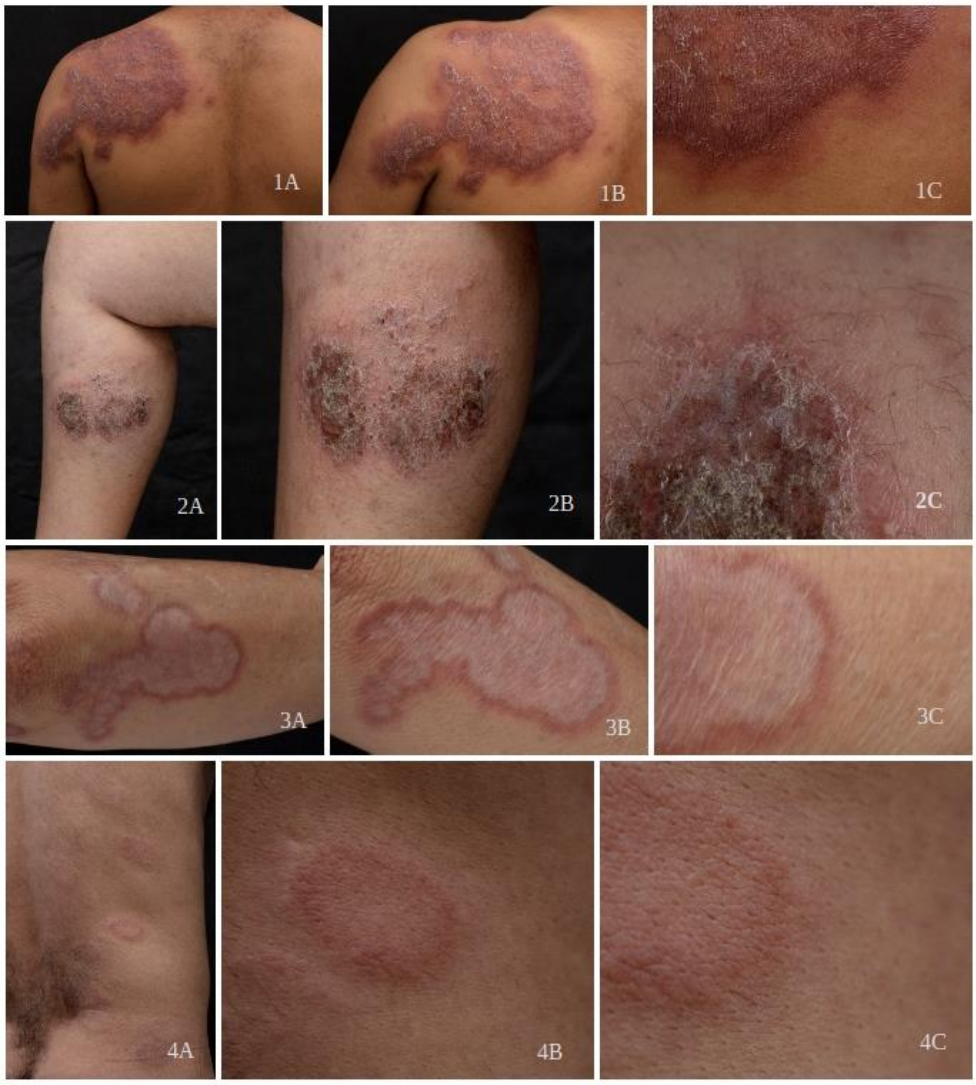
Supplementary Figure 1 – Artificial intelligence for leprosy study workflow from informed consent until data analysis.



Supplementary Figure 2 - Clinical aspects of paucibacillary forms of leprosy: A) indeterminate leprosy (I)- a single hypopigmented well defined macule on the left leg; B) tuberculoid leprosy (TT) skin lesion on the knee - plaque with infiltrated borders and central area with healthy aspect; C) borderline tuberculoid (BT)- a little plaque on the face with elevated borders; D) BT- hypopigmented plaque on the forearm with a slight erythema and infiltration; E) BT- a irregular plaque with erythematous papular edges; F) BT- a large hypopigmented, well defined and slightly infiltrated plaque on the arm; G) BT – erythematous-infiltrated plaque on the trunk with a smaller satellite lesion; H) BT - a large and infiltrated plaque with raised and scaly edge on the arm.



Supplementary Figure 3 - Clinical features of multibacillary leprosy cases. A) lepromatous leprosy (LL) case with multiples firm nodules (lepromas) of different sizes on the lower back; B) the same case showing lepromas spread throughout the trunk; C) Patient with lepromatous leprosy (LL) with severe saddle-nose deformity; D) LL patient exhibiting loss of eyebrow hair; E) LL case- small lepromas and diffuse infiltration of the ear; F) LL case - atrophy of the thenar and hypothenar muscles, trauma injuries of the skin and distal bone finger reabsorption resulting from loss of sensitivity on the hand; G) borderline lepromatous leprosy (BL) - symmetrically infiltrated plaques distributed on the trunk; H) Typical of borderline borderline leprosy (BB) - foveolar skin lesion in lower limb- a erythematous infiltrated plaque with poorly delineated outer borders and well-defined internal borders, large and irregular.



Supplementary figure 4- Examples of the imagens taken from each skin lesion: a panoramic photo, a close-up and the edge of the lesion including surrounding normal skin. 1A) Borderline lepromatous leprosy patient – a large erythematous-infiltrated plaque on the trunk - panoramic photo; 1B) close-up; 1C) photo from edge of the plaque; 2A) Other dermatological disease (ODD) – eczema plaque that break up into small papules and vesicles with crusts - panoramic photo; 2B) close-up; 2C) edge; 3A) ODD - granuloma annulare - erythematous plaque with well-defined borders, central pallor and atrophy; 3B) close-up photo; 3C) photo from the edge of the lesion; 4A) Leprosy – type 1 reaction- multiples erythematous-infiltrated plaques on trunk ; 4B) close-up; 4C) edge. Image storage followed the requirements with the same consultation room being used as a studio, with black background and floor (using a black backdrop). Images were photographed with artificial light sources 2 to 2.5 meters above the patient. Indirect intense lights (flashes) were combined with the continuous lights, while umbrellas using 500W tripod mounted moonlight (Portalite Elinchrom) supported the image taking. Images were recorded avoiding shadows, attaining maximum sharpness with texture, 3D and depth. The dermatologist selected the area to be photographed, aiming to have the lesion at the center (focus point) and to include surrounding healthy skin. The photographer used a Canon 5D high-resolution digital 50MP camera with 100 mm macro lenses (Canon Inc, Japan) with 1:1 magnification was used to avoid significant bias on background color. For the same reason, only one photographer was involved. The output was images of 50 x 70 cm with a 300dpi. Measurement with a meter scale were used in the camera during image acquisition and digital imaging software. The camera was color calibrated and white balanced according to manufacturer´s instructions.

Photos were never taken from areas that could identify the patient (whole face, tattoos, congenital lesions). When necessary, clothing, earrings, necklaces, or other identifiable objects were removed before photographing. Up to six lesions per patient and up to three images per lesion were captured. First a panoramic photo would be taken about 0.8 - 1.5m away from the lesion, to identify the body part. The second image was a close-up of the lesion (at about 10 to 15 cm distance, and when possible, a third image was taken from the edge of the lesion with some normal skin around it.

Supplementary Table 1

Feature	Leprosy (N = 100)	ODD (N = 122)	p-value
Age	51 (14 – 78)	48 (8 – 87)	0.024
Gender			
Male	66 (66.0%)	40 (32.8%)	<0.0001
Female	34 (34.0%)	82 (67.2%)	
Household contact			
Yes	29 (29.3%)	23 (19.0%)	0.081
No	70 (70.7%)	98 (81.0%)	
Duration of symptoms			
< 1	4 (4.0%)	12 (9.8%)	0.307
1 – 12	49 (49.0%)	63 (51.6%)	
12 – 24	17 (17.0%)	17 (13.9%)	
> 24	30 (30.0%)	30 (24.6%)	
Sensory loss *			
Present	18 (18.0%)	6 (4.9%)	0.0021
Absent	82 (82.0%)	116 (95.1%)	
Pruritus			
Present	7 (7.0%)	44 (36.1%)	<0.0001
Absent	93 (93.0%)	78 (63.9%)	
Pain			
Present	5 (5.0%)	10 (8.2%)	0.426
Absent	95 (95.0%)	112 (91.8%)	
Symptoms			
Present	28 (28.0%)	53 (43.4%)	0.025
Absent	72 (72.0%)	69 (56.6%)	
Site of lesion			
Face	35 (18.0%)	18 (10.6%)	0.025
Trunk	66 (34.0%)	48 (28.2%)	
Limbs	93 (47.9%)	104 (61.2%)	
Lesion type			
Macule	17 (14.7%)	34 (26.2%)	0.0002
Patch	67 (57.8%)	86 (66.2%)	
Papule	5 (4.3%)	3 (2.3%)	
Nodule	27 (23.3%)	7 (5.4%)	
Vesicle			
Present	0 (0.0%)	3 (2.5%)	0.25
Absent	100 (100.0%)	119 (97.5%)	
Blister			
Present	1 (1.0%)	0 (0.0%)	0.45
Absent	99 (99.0%)	122 (100.0%)	
Loss of eyebrows or eyelashes			
Yes	14 (14.0%)	0 (0.0%)	<0.0001
No	86 (86.0%)	122 (100.0%)	

Cont.

Feature	Leprosy (N = 100)	ODD (N = 122)	p-value
Diffuse infiltration			
Present	18 (18.0%)	0 (0.0%)	<0.0001
Absent	82 (82.0%)	122 (100.0%)	
Paresthesis in hands			
Present	16 (16.0%)	3 (2.5%)	0.0004
Absent	84 (84.0%)	119 (97.5%)	
Paresthesis in feet			
Present	30 (30.0%)	1 (0.8%)	<0.0001
Absent	70 (70.0%)	121 (99.2%)	
Thermal sensitivity loss[a]			
Present	69 (69.0%)	25 (20.5%)	<0.0001
Absent	31 (31.0%)	97 (79.5%)	
Pain sensitivity loss[b]			
Present	16 (16.0%)	4 (3.3%)	0.002
Absent	84 (84.0%)	118 (96.7%)	
Tactile sensitivity loss[c]			
Present	3 (3.0%)	0 (0.0%)	0.09
Absent	97 (97.0%)	122 (100%)	
Lichenification			
Present	0 (0.0%)	15 (12.3%)	0.0002
Absent	100 (100.0%)	107 (87.7%)	
Hyperesthesia			
Present	6 (6.0%)	0 (0.0%)	0.008
Absent	94 (94.0%)	122 (100.0%)	
Scaling surface			
Present	5 (5.0%)	37 (30.3%)	<0.0001
Absent	95 (95.0%)	85 (69.7%)	
Ulceration			
Present	1 (1.0%)	5 (4.1%)	0.23
Absent	99 (99.0%)	117 (95.9%)	
Enlarged nerve			
Present	8 (8.0%)	0 (0.0%)	0.0014
Absent	92 (92.0%)	122 (100.0%)	
Number of lesions			
1	15 (15.0%)	31 (25.4%)	<0.0001
2 – 5	14 (14.0%)	28 (23.0%)	
6 – 10	2 (2.0%)	14 (11.5%)	
11 – 20	11 (11.0%)	13 (10.7%)	
> 20	58 (58.0%)	36 (29.5%)	

Cont.

Feature	Leprosy (N = 100)	ODD (N = 122)	p-value
Ridley-Jopling classification			
TT	3 (3.3%)		
BT	24 (26.4%)		
BB	12 (13.2%)		
BL	19 (20.9%)		
LL	28 (30.8%)		
Indeterminate	5 (5.5%)		

Supplementary Table 1- Clinical and demographical features from patients with leprosy and other dermatological diseases.

*As defined by the dermatologist, a- Thermal sensitivity loss (tubes containing warm and cold water are used), b-pain sensitivity loss (a sharp stick or needle to assess the sensation of pain), c- tactile sensitivity loss (assessed using a piece of cotton).

Supplementary Table 2

	Odds Ratio	95% Confidence Interval
Lesion type (papule & nodule vs. macule)	36.97	(11.89,76.67)
Lesion color (hyporpigmented vs. erythematous)	0.55	(0.25,1)
Lesion color (hyperpigmented vs. erythematous)	0.41	(0.22,0.85)
Site of lesion (trunk vs. face)	1.32	(1,2.08)
Site of lesion (lower limbs vs. face)	0.56	(0.35,0.95)
Sensory loss (present vs. absent)*	5.26	(1.16,24.53)
Thermal sensitivity loss (present vs. absent)	3.16	(1,13.04)
Tactile sensory loss (present vs. absent)	2.74	(1,7.73)
Pruritus (present vs. absent)	0.80	(0.35,1)
Hyperaesthesia (present vs. absent)	37.71	(6.98,95.48)
asymptomatic (present vs. absent)	4.22	(1.65,7.45)

Supplementary table 2. Model 2 by elastic-net logistic regression. We used elastic-net logistic regression with repeated 10-fold cross validation on the complete dataset including the subset features to get the final model. Bootstrapping was used to construct the 95% confidence interval.

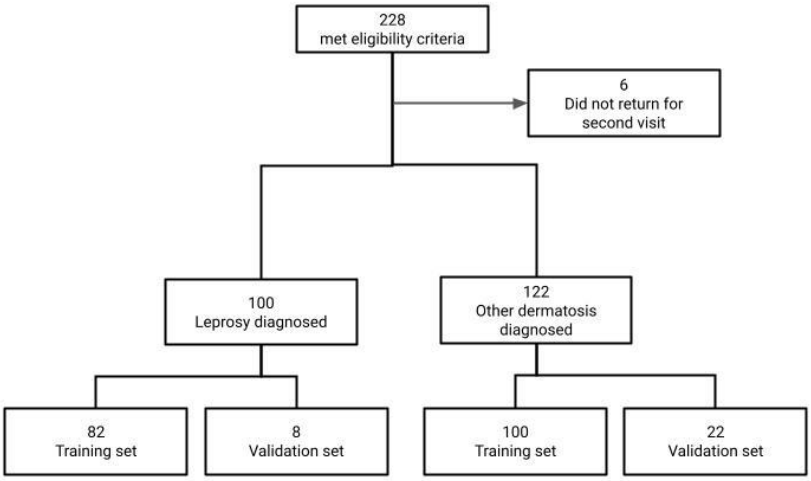
*as measured and defined by the dermatologist.

Supplementary Table 3

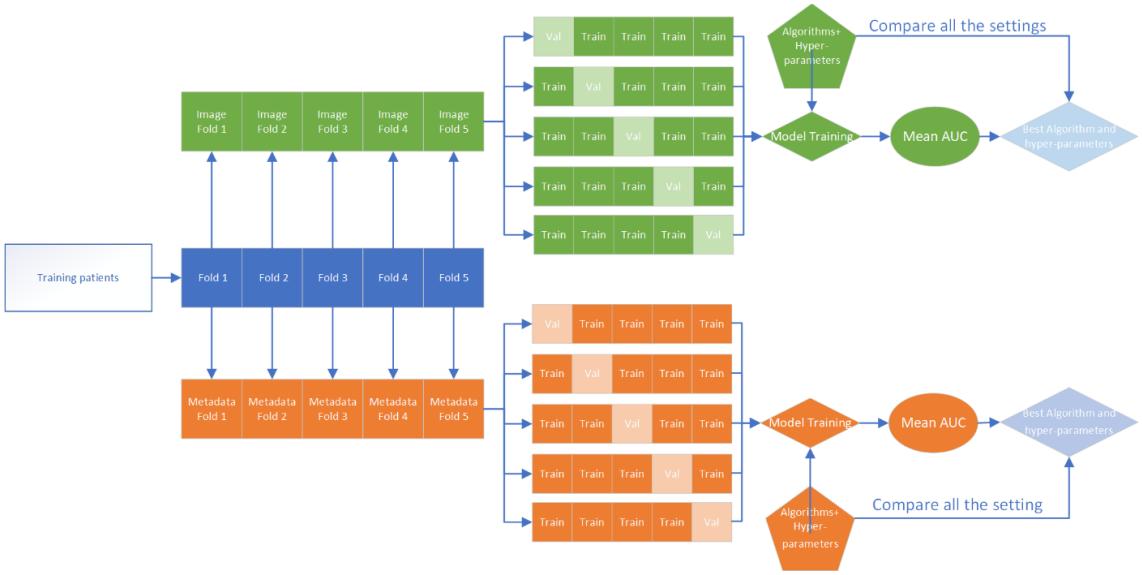
	Odds Ratio	95% Confidence Interval
Minimum Model 2 Probability	2.74	(1,44.88)
Median Model 2 Probability	2.74	(1,20.25)
Mean Model 2 Probability	2.91	(1,19.68)
Max Model 2 Probability	2.8	(1,23.56)
Ratio of Model 2 Probabilities over 50%	1.79	(1,6.91)
Gender (male vs female)	1.9	(1.35,7.36)
Sensory loss (present vs absent)*	1.02	(1,2.45)
Pain (present vs absent)	0.93	(0.2,1)
Face lesion (present vs absent)	1.17	(1,3.41)
Trunk lesion (present vs absent)	1.15	(1,2.56)
Arm/leg lesion (present vs absent)	1.16	(1,3.98)
Loss of eyebrows or eyelashes (yes or no)	1.26	(1,2.62)
Diffuse infiltration (present vs absent)	1.48	(1,4)
Paresthesis in feet (present vs absent)	1.93	(1,21.5)
Thermal sensitivity loss (present vs absent)	1.21	(1,1.96)
Lichenification (present vs absent)	0.78	(0.2,1)
Scaling surface (present vs absent)	0.74	(0.2,1)
Number of lesions >20 (yes or no)	1.48	(1,4.47)

Supplementary Table 3. Model 3 by elastic-net logistic regression using Model 2 outputs and patient information. We used elastic-net logistic regression with repeated 10-fold cross validation on the complete dataset including the testing patients to get the final model. Bootstrapping was used to construct the 95% confidence interval.

*patient information



Supplementary figure 5- Flow STROBE diagram to present enrolled patients in the study.



Supplementary Figure 6- Learning algorithm selection and hyperparameter tuning for Model 1 and Model 2. Data partitioning for experiments for a validation step. A fifth of the patients was separated for testing, while data from the remaining patients were used to train the three machine learning models. Tuning parameters were selected by 10-fold cross validation repeated for LR and XGB and by out-of-bag errors for RF. Before training the models, we first partitioned all the 182 patients into five folds, as shown in figure below. Then, we used 5-fold cross-validation to evaluate the performance for both Model 1 and 2, which is widely used to assess how a model will generalize to unseen data. In each trial, there were always: the inner training group including lesion images or metadata from four folds of the patients and the inner testing group including the rest of data. A model was trained by a learning algorithm using the inner training group and then validated on the inner testing group. We repeated this procedure for five times until each of the five folds had been used exactly once as the inner testing group. All the metrics in Table 2 and 3 are averages of the five trials. Note that the partition is the same for all the experiments for Model 1 and 2. Thus, we could make a fair comparison of all the learning algorithms.