Received: 3 November 2020 | Accepted: 2 December 2020 | Published online: 16 February 2021

DOI: 10.1002/dad2.12146

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

SHORT REPORT

Deep learning shows no morphological abnormalities in neutrophils in Alzheimer's disease

Floris Chabrun[1,2] | **Xavier Dieu**[1,2] | **Nicolas Doudeau**[1] | **Jennifer Gautier**[3] | **Damien Luque-Paz**[2,4] | **Franck Geneviève**[4] | **Marc Ferré**[2] | **Delphine Mirebeau-Prunier**[1,2] | **Cédric Annweiler**[2,3] | **Pascal Reynier**[1,2]

[1] Département de Biochimie et Biologie Moléculaire, Centre Hospitalier Universitaire d'Angers, Angers, France

[2] Unité Mixte de Recherche MITOVASC, CNRS 6015/INSERM 1083, Université d'Angers, Angers, France

[3] Département de Gériatrie, Centre Hospitalier Universitaire d'Angers, Angers, France

[4] Département d'Hématologie, Centre Hospitalier Universitaire d'Angers, Angers, France

Correspondence
Floris Chabrun, Centre Hospitalier Universitaire d'Angers, 4 rue Larrey, 49933 Angers Cedex 9, France.
E-mail: floris.chabrun@chu-angers.fr

Funding information
University of Angers and the University Hospital of Angers

Abstract

Introduction: Several studies have provided evidence of the key role of neutrophils in the pathophysiology of Alzheimer's disease (AD). Yet, no study to date has investigated the potential link between AD and morphologically abnormal neutrophils on blood smears.

Methods: Due to the complexity and subjectivity of the task by human analysis, deep learning models were trained to predict AD from neutrophil images. Control models were trained for a known feasible task (leukocyte subtype classification) and for detecting potential biases of overfitting (patient prediction).

Results: Deep learning models achieved state-of-the-art results for leukocyte subtype classification but could not accurately predict AD.

Discussion: We found no evidence of morphological abnormalities of neutrophils in AD. Our results show that a solid deep learning pipeline with positive and bias control models with visualization techniques are helpful to support deep learning model results.

KEYWORDS
Alzheimer's disease, artificial intelligence, deep learning, machine learning, neutrophils

1 | INTRODUCTION

Over the past decade, inflammation has emerged as a prominent feature of Alzheimer's disease (AD) pathophysiology.[1] Several studies have provided evidence of a key role played by neutrophils in AD-related inflammatory processes,[2] mediated by intracellular granules, surface expression of inflammatory proteins, higher extravasation, and vascular NETosis.[2–4]

It is known that, on a blood smear, neutrophils exhibit abnormalities or morphologic particularities during infections and inflammatory processes.[2,5,6] However, to our knowledge, no studies have investigated such potential morphological particularities of AD neutrophils, secondary to their functional alteration. The objective of the present study is to investigate the presence of such abnormalities on peripheral blood smears and their potential as a biomarker of diagnosis of AD.

Detecting unknown abnormalities on a blood smear is a complex and time-consuming task. Deep learning has emerged during the past decade as a powerful tool for image analysis, capable of achieving expert-level accuracy and even discovering unknown image patterns of clinical importance.[7–9] To explore the link between neutrophil

morphology and AD, we trained state-of-the-art artificial intelligence architectures based on machine learning methods, so-called deep learning, to predict AD based on images of blood smears.

2 | MATERIALS AND METHODS

2.1 | Cohort construction and data collection

Patients from the Department of Geriatrics of the University Hospital of Angers were recruited into two cohorts: the AD group for patients with AD and the subjective memory complaint (SMC) group for patients without dementia and without mild cognitive impairment (MCI) who had normal neuropsychological and functional performance. Details concerning the recruitment of patients are listed in File S1, section 1: Cohort construction in supporting information. Leukocyte images for patients were retrieved from routine a health-care database of analyses run in the Department of Hematology of the University Hospital of Angers. Those images were partitioned in training, validation, and test sets according to rules described in File S1, section 2: Data partitioning.

This study was performed using anonymized data collected during the care and treatment of patients. According to the approval of the Ethics Committee of the University Hospital of Angers (no 2020/118), due to the tacit consent of patients for the use of their anonymized data collected during routine care, written consent was not required and replaced by verification of the absence of patient's opposition. Data were collected and processed after declaration to the University Hospital of Angers' Data Protection Officer according to the General Data Protection Regulation (declaration ar19-0046v0).

2.2 | Prediction tasks

Two pre-trained state-of-the-art architectures (VGG-16 and Inception v3) were trained after transfer learning for three learning prediction tasks, including two control tasks, to ensure the veracity of the results obtained for AD patients.

The first task, referred to as "AD prediction," aimed at classifying patients with AD versus patients with SMC from neutrophil images from peripheral blood smears.

The same deep learning architectures were used for a simpler task: the prediction of already known imaging features, namely leukocyte subtype classification into four classes: neutrophils (NE), eosinophils (EO), monocytes (MO), and lymphocytes (LY). Basophils were excluded because of an insufficient number of samples.

Finally, to explore biases potentially responsible for overfitting, by learning by heart for example, we also trained the same architectures to predict the patient to whom each image belonged. This task is referred to as "patient identification."

Details concerning the choice of those deep learning architectures and the training are presented in File S1, sections 3-7; see Figure S1 in supporting information.

HIGHLIGHTS

- There is no argument in favor of morphologically abnormal neutrophils in Alzheimer's disease.
- Study is based on data allowing state-of-the-art leukocyte deep learning classification.
- A pipeline allows its robustness to be checked regardless of sample size.

RESEARCH IN CONTEXT

1. **Systematic review**: We reviewed literature using published and pre-print sources (eg, PubMed, arXiv). Several studies have recently provided evidence of a major role played by neutrophils in the pathophysiology of Alzheimer's disease. To our knowledge, however, no study has addressed the presence of morphological alterations among neutrophils on blood smears, despite the well-known links between inflammation processes and neutrophilic morphological abnormalities.
2. **Interpretation**: Our results confirm the absence of neutrophilic morphological abnormalities on blood smears. This is supported by deep learning control models achieving state-of-the-art accuracy for a known feasible task on the same data and visualization techniques exploring the behavior of those models.
3. **Future directions**: As neutrophils did not show any morphological alterations on blood smears carried out on routine blood formula counts, it may be interesting to analyze higher resolution images to search for finer alterations or focus on non-cell elements not routinely analyzed such as neutrophil extracellular traps.

3 | RESULTS

3.1 | Cohort description

The cohort is described in Table 1. Two patients had Waldenström's disease: one in the AD group and one in the SMC group. Because of the balance between both groups, those patients were kept in the analyses. No other patients in the AD group had any form of malignant hemopathy. In the SMC group, one patient had chronic lymphocytic leukemia, and one had a follicular lymphoma. Those patients were kept in our study, because our partitioning pipeline prevents models from being tested on images of patients they were trained with.

TABLE 1 Cohort description

Group	AD	SMC	Total/p-value[2]
Patients			
Number of patients	8	10	Total: 18
Number of reports	12	13	Total: 25
Age: mean ± SD	85.3 ± 5.4 years	84.8 ± 6.6 years	N.S.[1]
Biology			
Total leukocytes count	10.8 G/L	8.0 G/L	N.S.[1]
PMN count	8.6 G/L	6.6 G/L	N.S.[1]
PME count	0.2 G/L	0.1 G/L	* 1
Monocyte count	0.7 G/L	0.5 G/L	N.S.[1]
Lymphocyte count	1.2 G/L	0.6 G/L	* 1
Hemoglobin	11.2 g/dL	11.8 g/dL	N.S.[1]
Hematocrit	33.0%	35.2%	N.S.[1]
Mean globular volume	92.5 fL	96.3 fL	N.S.[1]
Mean hemoglobin concentration	33.8	33.7	N.S.[1]
Platelet count	211	152	N.S.[1]
CRP ≥ 5 mg/L/total	7/9 (3 unknown)	7/8 (5 unknown)	N.S.[2]
Images			
PMN images	1455	1468	Total: 2923
PME images	54	28	Total: 82
Monocyte images	117	118	Total: 235
Lymphocyte images	347	200	Total: 547

Notes: *P* values were computed using Mann-Whitney tests. N.S.: $P > 0.05$.
*$P < 0.05$.
[1]*P* values were determined with Mann-Whitney tests.
[2]*P* value was determined with a Fisher's exact test.
Abbreviations: CRP, C-reactive protein; PME, polymorphonuclear eosinophil; PMN, polymorphonuclear neutrophil; SD, standard deviation.

3.2 | Leukocyte subtype classification

The top model (VGG-16) achieved an accuracy on the test set of 97.5% (94.3% with random translation/rotation applied to the image). For all models, median accuracy was 84.6% (96.1% and 82.2% median accuracy for models trained with soft and strong image augmentation, respectively).

3.3 | Alzheimer's disease prediction

The top model (VGG-16) achieved an area under the curve for receiver operating characteristic curve (AUC-ROC) on the test set of 0.68 (0.62 with random translation/rotation applied to the image). For all models, the median AUC-ROC was 0.5.

3.4 | Patient identification

The top model (Inception v3) achieved an accuracy on the test set of 95.3% (80.3% with random translation/rotation applied to the image). For all models, the median accuracy was 17.3% (73.8% and 12.2% median accuracy for models trained with soft and strong image augmentation, respectively).

3.5 | Gradient-weighted class activation mapping (Grad-CAM) visualization

The heatmaps presented in Figure 1A–C depict the attention of neural networks on the image for the associated output prediction. This allows assessment of (1) which regions are useful for the predictions

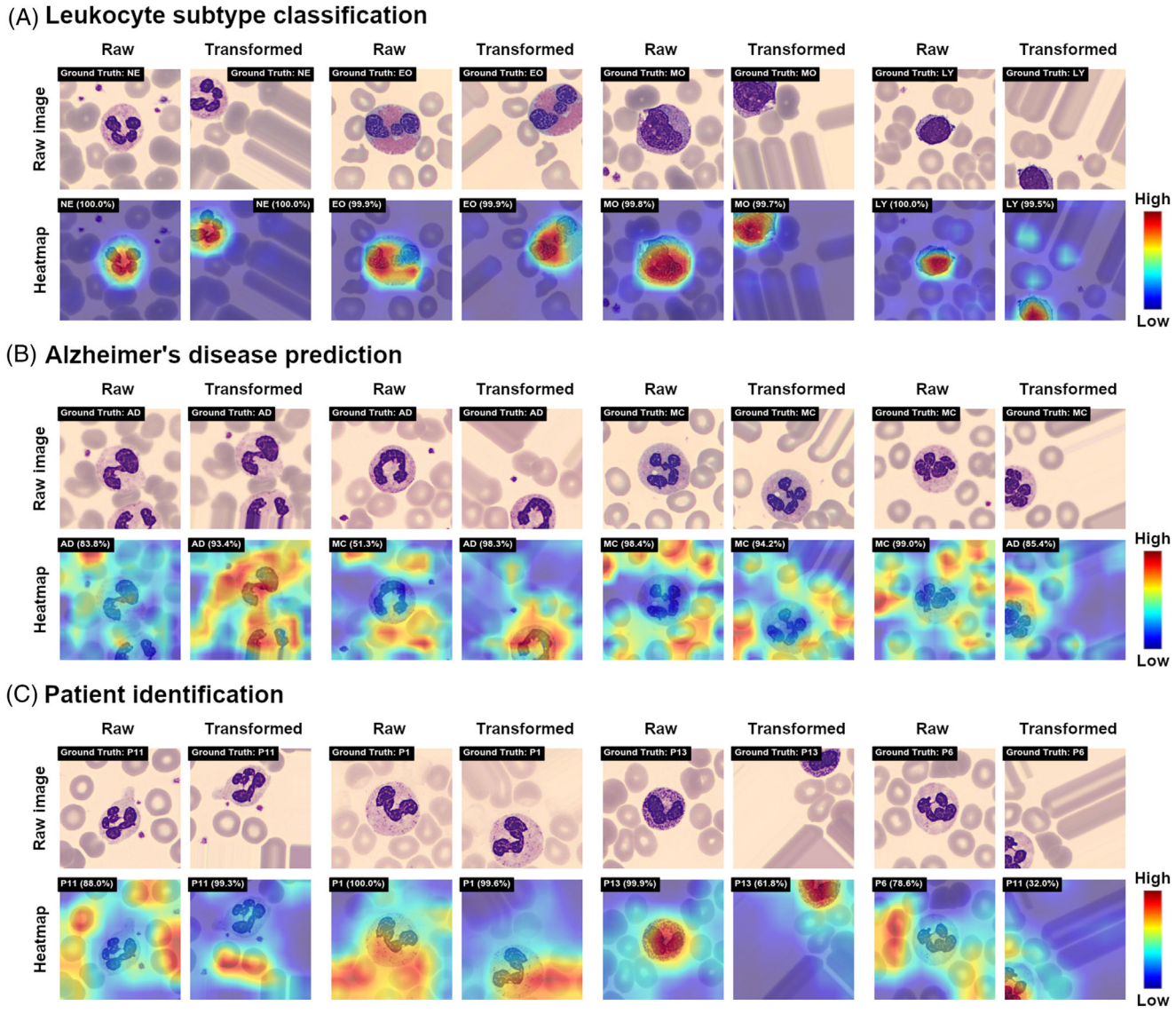

FIGURE 1 Gradient-weighted class activation mapping (Grad-CAM) visualization of best models for blood cell type classification (A), Alzheimer's disease prediction (B), and patient prediction (C). For each image, the model was inputted with the raw image (left) and the same image after random translation and/or rotation (right). Predicted class and associated confidence are plotted on the heatmap, while ground truth is plotted on the input image

and whether those regions are consistent with the expectations of human hematologists, namely for leukocyte subtype prediction; (2) the number and size of those regions; and (3) the stability and robustness of neural networks when transformations such as translation or rotation are applied to the image.

4 | DISCUSSION

With 18 patients and 3787 pictures, we trained a model to classify cells according to the four main leukocyte types, with an overall state-of-the-art accuracy of 97.3%.[10–12] The consistency of those results was confirmed by (1) Grad-CAM, showing that the attention of the model was focused on the area an expert hematologist would interpret, that is, cells' nuclei and cytoplasm, and this attention follows the elements when translation/rotation is applied to the image; and (2) the results during grid search, with most models achieving an overall accuracy on the validation set higher than 85%, even with strong image augmentation.

We also showed that the deep learning models used (mainly Inception v3) could achieve a high accuracy when predicting to which patient a cell image belongs. Grad-CAM showed that this ability is attained by focusing on more minor details, such as red blood cell density and spreading, or particularities on cells, like granulation density. Although this highlights the high sensitivity of those small details and thus their possible performance, this also highlights how those models are prone to overfitting. This is demonstrated by predictions that neural networks tend to change when translation/rotation is applied to the image.

The best model trained for AD prediction showed an interesting AUC of 0.70 on the training set, confirmed on a test set of new images (AUC = 0.68). However, Grad-CAM visualization supports a tendency to overfit, showing a high versatility concerning the details on which the model focuses and a high instability when applying translation or rotation to the images. Furthermore, a grid search showed that most models could not achieve an AUC on a test set higher than 0.5, thus supporting this tendency.

Our results draw attention to the overfitting risks of deep learning architectures. We hereby show that state-of-the-art deep learning architectures can efficiently learn to memorize to which patient an image belongs, based on a variety of details, some of which are minute. In our study, we carefully partitioned our samples to make sure that a model trained on images from a patient would not be tested on images of the same patient, to avoid any bias leading to an overestimation of its performance. However, to date, a number of articles recently applying deep learning to blood cell classification or prediction have not specified the stratification during partitioning.[12–14] This omission may lead to the overestimation of both the results of those models and particularly their ability to generalize to new patients.

4.1 | Limits

Our study was restricted on retrospective data analyzed during the routine care of patients seen in the University Hospital of Angers. For further confirmation of these results, multicentric prospective data, including blood smear images obtained with other coloration methods and other digital microscopes will be needed. Our data are accessible for further analysis upon reasonable request to the corresponding author.

5 | CONCLUSIONS

We used a solid prediction-based pipeline: (1) a positive control predictive model, trained to verify the feasibility of a known task on our dataset, that is, leukocyte subtype classification; (2) a bias control predictive model, trained to determine whether models could learn features allowing them to overfit; and (3) Grad-CAM visualization to check that the attention given by the model fit with our expectations. Thanks to this pipeline, we have demonstrated here that there is no evidence of a link between blood smear neutrophil morphology and AD.

We have also highlighted the importance of control predictive models, heavy visualization techniques, grid search, and transfer learning, enabling us to verify the consistency of models and prevent falsely accurate results from being trusted, despite a complex task based on a relatively low number of samples.

ACKNOWLEDGMENTS

This work was granted access to the HPC resources of IDRIS under the allocation 2019-AD011011303 made by GENCI.

AUTHORS' CONTRIBUTION

FC: design of the study and the deep learning methodology, data processing, statistical and deep learning analyses, literature review, and redaction of the manuscript; XD: design of the deep learning methodology, review of statistical and deep learning analyses; **ND**: literature review concerning the deep learning methodology, data processing, and deep learning analysis; **JG**: collection and pre-processing of clinical data; DL: collection of biological data and review of hematology-related sections and literature; FG: collection of biological data and review of hematology-related sections and literature; MF: collection of biological data and review of the deep learning methodology; DM: global review of the methodology and proofreading of the manuscript; CA: design of the study, collection and structuring of clinical data, review of Alzheimer's disease-related sections and literature, global review of the methodology, and proofreading of the manuscript; PR: design of the study, global review of the methodology, and proofreading of the manuscript.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest.

ORCID

Floris Chabrun https://orcid.org/0000-0001-6871-8711

REFERENCES

1. Kinney JW, Bemiller SM, Murtishaw AS, Leisgang AM, Salazar AM, Lamb BT. Inflammation as a central mechanism in Alzheimer's disease. *Alzheimers Dement (N Y)*. 2018;4:575-590. https://doi.org/10.1016/j.trci.2018.06.014.
2. Dong Y, Lagarde J, Xicota L, et al. Neutrophil hyperactivation correlates with Alzheimer's disease progression. *Ann Neurol*. 2018;83:387-405. https://doi.org/10.1002/ana.25159.
3. Pietronigro E, Zenaro E, Constantin G. Imaging of Leukocyte Trafficking in Alzheimer's. *Disease Front Immunol*. 2016;7:33. https://doi.org/10.3389/fimmu.2016.00033.
4. Stock AJ, Kasus-Jacobi A, Pereira HA. The role of neutrophil granule proteins in neuroinflammation and Alzheimer's disease. *J Neuroinflammation*. 2018;15:240. https://doi.org/10.1186/s12974-018-1284-4.
5. Karakonstantis S, Koulouridi M, Pitsillos K, et al. A prospective study of hospitalized patients with leukemoid reaction; causes, prognosis and value of manual peripheral smear review. *Rom J Intern Med*. 2019;57:241-247. https://doi.org/10.2478/rjim-2019-0006.
6. Zini G, Bellesi S, Ramundo F, Onofrio G. Morphological anomalies of circulating blood cells in COVID-19. *Am J Hematol*. 2020;95:870-872. https://doi.org/10.1002/ajh.25824.
7. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical Image Analysis using Convolutional Neural Networks: a Review. *J Med Syst*. 2018;42:226. https://doi.org/10.1007/s10916-018-1088-1.
8. Matek C, Schwarz S, Spiekermann K, Marr C. Human-Level Recognition of Blast Cells in Acute Myeloid Leukaemia with Convolutional Neural Networks. *Nature Machine Intelligence*. 2019;1(11):538–544. https://doi.org/10.1038/s42256-019-0101-9.
9. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. 2018;2:158. https://doi.org/10.1038/s41551-018-0195-0.

10. Acevedo A, Alférez S, Merino A, Puigví L, Rodellar J. Recognition of peripheral blood cell images using convolutional neural networks. *Comput Methods Programs Biomed*. 2019;180:105020. https://doi.org/10.1016/j.cmpb.2019.105020.
11. Habibzadeh M, Jannesari M, Rezaei Z, Baharvand H, Totonchi M. Automatic white blood cell classification using pre-trained deep learning models: resNet and Inception. *Tenth International Conference on Machine Vision (ICMV 2017)*. International Society for Optics and Photonics; 2018:1069612. https://doi.org/10.1117/12.2311282.
12. Yu W, Chang J, Yang C, et al. Automatic classification of leukocytes using deep neural network. *2017 IEEE 12th International Conference on ASIC (ASICON)*. Guiyang: IEEE; 2017:1041-1044. https://doi.org/10.1109/ASICON.2017.8252657.
13. Qin F, Gao N, Peng Y, Wu Z, Shen S, Grudtsin A. Fine-grained leukocyte classification with deep residual learning for microscopic images. *Comput Methods Programs Biomed*. 2018;162:243-252. https://doi.org/10.1016/j.cmpb.2018.05.024.
14. Wang Q, Bi S, Sun M, Wang Y, Wang D, Yang S. Deep learning approach to peripheral leukocyte recognition. *PLoS One*. 2019;14:e0218808. https://doi.org/10.1371/journal.pone.0218808.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Chabrun F, Dieu X, Doudeau N, et al. Deep learning shows no morphological abnormalities in neutrophils in Alzheimer's disease. *Alzheimer's Dement*. 2021;13:e12146. https://doi.org/10.1002/dad2.12146