



OPEN

A literature-derived dataset on risk factors for dry eye disease

DATA DESCRIPTOR

Wanju Yang¹, Kuiliang Yang², Yumiao Pan¹, Shangcao Wu¹, Xiangxi Chen¹, Lei Shen¹, Qingyan Zeng¹, Jianhua Wu¹, Minzhi Lv^{3,4}, Junyan Zhang⁵, Yanning Yang² & Yiqiao Xing¹

Dry eye disease (DED) is a common disease associated with disorder of tear secretion. Research on risk factors for DED, such as depression, arthritis, thyroid disease, stroke and diabetes, is important to facilitate its diagnosis and prognosis. We created a dataset on risk factors for DED (DrDED) with public access that can provide up-to-date and validated data acquired from systematically searched and screened, high-quality studies. The established DrDED contained 119 studies published between 2000 and 2022. The range of the study sample size was from 43 to 4,871,504. The study types were, as follows: cross-sectional ($n=92$), retrospective cohort ($n=9$), prospective cohort ($n=10$), and case-control ($n=8$) studies. Data from eligible studies were collected and presented for the present study, including the publication information, study characteristics, definition and prevalence of the disease, and risk factors for DED, together with the strength of association. With the publication of new relevant studies, the DrDED will be updated, and the data will be made accessible to the users.

Background & Summary

Dry eye disease (DED) is an inflammatory and chronic clinical condition, and characteristic for reduced tear film stability that affects the ocular surface^{1,2}. DED occurs when any of the three layers of a healthy tear film (aqueous fluid, mucus, and fatty oils) is affected, and as a result, the surface of the eyes is inadequately lubricated, and became less smooth and clear³.

The prevalence of DED is high among ophthalmic conditions, and varies between 14.6% and 30%⁴, depending on the diagnostic method used and population characteristics⁵. A recent systematic review revealed that one of five Asians experience DED⁶. Hence, DED has been considered a disease of multifactorial etiology, and is correlated to the presence of certain comorbidities in systemic organs⁷. DED etiology and pathology findings have indicated that patients with different DED etiologies may present with different signs and symptoms⁸. It has been reported that patients who smoke or have comorbidities experience more severe symptoms of DED, when compared to patients in a control group⁹. The potential factors that might increase the incidence of DED include primary lacrimal gland diseases, rheumatoid arthritis, ultraviolet ray exposure, improper wearing of contact lenses, older age, female, smoking, and some medicines, such as tricyclic antidepressants and antihypertensive drugs^{3,6,10–12}.

The categories of risk factors for DED can be classified as modifiable and non-modifiable, and further classified into three categories: consistent, probable and inconclusive¹³. For example, a modifiable factor, alcohol drinking, was considered inconclusive, in terms of its association with DED. On the contrary, Asian ethnicity is considered an essential, non-modifiable and consistent risk factor for DED¹⁴. The observed increase in incidence of DED in recent years may be associated with the growing exposure to risk factors in the general population, such as increased sunlight exposure, smoking, and aging^{12,15}. The examination of DED risk factors and the strength of its association with DED can provide valuable information for the prevention and treatment of DED.

The present study established a dataset on risk factors for dry eye disease (DrDED¹⁶), which comprised of published literatures of high-quality clinical studies that reported the DED risk factors and its associative strengths. At present, the DrDED contains the following data obtained from 119 studies: publication details, geographic regions from where the study subjects were obtained, number and demographic characteristics of the study subjects, the DED risk factors under investigation, and the magnitude of the association strength with the disease in the form of statistical measures, including odds ratios (ORs) and confidence intervals (CIs).

¹Aier Eye Hospital of Wuhan University, Wuhan, Hubei, 430063, China. ²Eye Center, Renmin Hospital of Wuhan University, Wuhan, Hubei, 430060, China. ³Department of Biostatistics, Zhongshan Hospital, Fudan University, Shanghai, 200032, China. ⁴Center of Evidence-Based Medicine, Fudan University, Shanghai, 200433, China. ⁵Department of Clinical Study Designing, Bothwin Clinical Study Consultant, Redmond, WA, 98053, USA. ✉e-mail: ophyyn@gmail.com; Yiqiao_xing57@whu.edu.cn

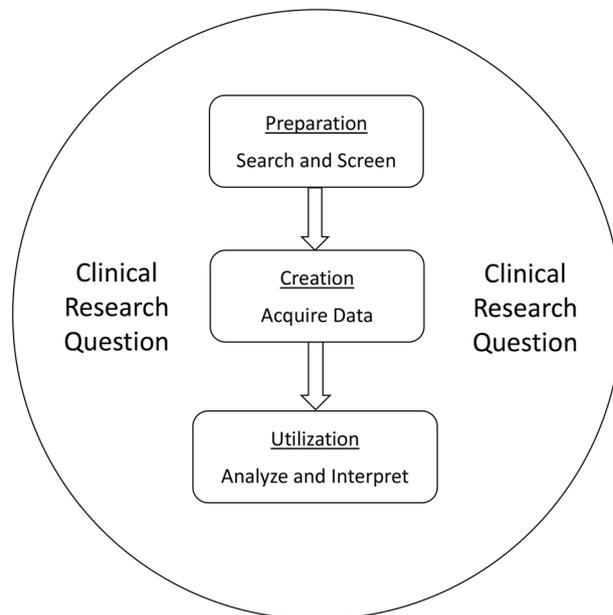


Fig. 1 Diagram for the development of the DrDED. DrDED, dataset on risk factors for dry eye disease.

Methods

Overview and literature searches. The present study was conducted based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines¹⁰. The clinical research question was initially formulated: what are the risk factors of DED and the strength of its association with the disease? The preparation stage for the DrDED development included a systematic literature search, the establishment of the inclusion and exclusion criteria, and the study screening through the evaluation of the eligibility of each study. Next, the DrDED was created by acquiring data correlated to the research question from eligible studies, and entering these data in a Microsoft Excel spreadsheet. After the data validation, the dataset was ready for utilization by researchers who had access to the DrDED. All stages were conducted with focus on the clinical research question (Fig. 1).

A systematic search strategy was performed using the following Medical Subject Headings terms and free-text terms: “dry eye syndrome”, “dry eye disease”, “dry eye”, “evaporative dry eye disease or syndrome”, “population at risk”, “risk factor”, “risk score”, and “health correlate” (the relationship among disease names and among risk factor-related terms, “OR”; the relationship between the disease and risk factors). The Medline (since 1946), Embase (since 1974), and Cochrane Central Register of Controlled Trials (since 1996) electronic databases were also searched from inception to September 2022. No restriction was made on the type of study. In order to ensure that no eligible study was missed, the references of previously published systematic reviews¹¹ and included studies were checked.

Inclusion and exclusion criteria. The following inclusion criteria were applied for the screening of the titles, abstracts and full-text articles: any type of primary patient research on risk factors or predictors for DED, including cross-sectional studies, case-control studies and cohort studies; DED was diagnosed according to the United States National Eye Institute/Industry Workshop criteria¹²; the data for risk factors in the DED analysis was provided; the appropriate methods for the study quality control were reported; a between-group comparability analysis was performed for the baseline data; confounding variables were adjusted in the calculation of ORs for the studied factors; complete study and participant information (author/s, year of publication, country, participant age, and gender) was available. Articles that reported duplicate data from the same study were excluded.

Risk of bias assessment. Two independent researchers performed the risk of bias assessment for each included study using the risk of bias tool, the Newcastle-Ottawa Scale (NOS), as recommended by the Cochrane Handbook for Systematic Reviews. The NOS consisted of eight items in three domains (selection, comparability and outcomes), and the total score ranged from 0 to 9. The studies were rated, as follows: low risk of bias (high quality study), when the NOS score ranged within 7–9; high risk of bias, when the NOS score ranged within 4–6; very high risk of bias (very low quality study), when the NOS score ranged within 0–3¹⁷.

Data Records

As presented in Fig. 2, among the 3,300 citations found in the extensive search, 210 articles were identified after screening the titles and abstracts, and after reviewing the full texts. A total of 119 relevant articles published between 1997 and 2022 met the study inclusion criteria¹⁶. The types of studies were, as follows: cross-sectional ($n = 92$), case-control ($n = 8$), retrospective cohort ($n = 9$), and prospective cohort ($n = 10$) studies. The range of the study sample size was from 43 to 4,871,504. The reported risk factors and strength of association with the

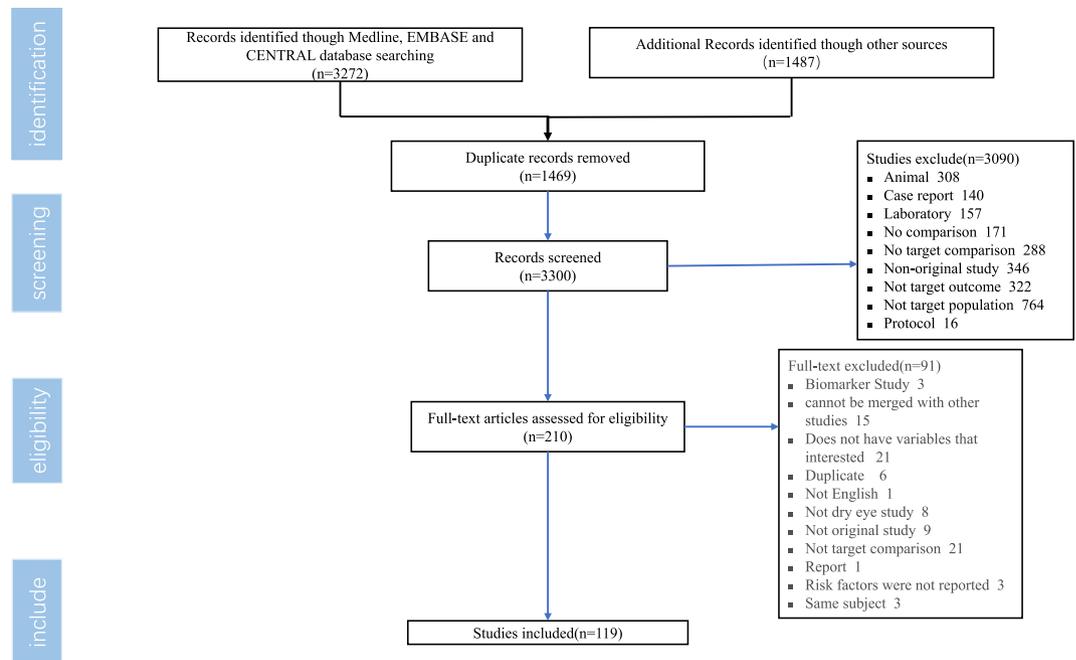


Fig. 2 Study eligibility assessment flow diagram for the DrDED. DrDED, dataset on risk factors for dry eye disease.

disease (measured by multiple logistic regression models and reported with ORs and 95% CIs) were acquired from eligible studies, and studies included in the DrDED¹⁶. Table 1 presents the structure of the DrDED. The data obtained from all eligible studies, according to the clinical research question, was entered in the database¹⁶.

A total of 34 risk factors were reported in the DrDED including thyroid disease, arthritis, hypertension, diabetes, stroke disease, smoking status, alcohol drinking status, depression, age (every 10 years, as a continuous variable for continuous variables or categorical variables as reported in the original studies), gender, caffeine use, antihistamines use, visual display terminal use, contact lens use, body mass index, site/geographical region, gout history, dyslipidemia, multivitamin use, pterygium, occupation or work activity environment, allergy, hormone replacement therapy, corticosteroid use, cataract or cataract surgery, glaucoma, residence at high altitude, under-correction of refractive error/refractive surgery, residence (rural or urban), benign prostatic hyperplasia, rosacea, diuretics, migraine headache, and sleep disorders.

In order to illustrate the utilizations of the DrDED, a subset of 41 representative studies published between January 2000 and April 2021, which reported one or more of the eight common risk factors for DED, was selected. In these studies, the participants were recruited from the same geographic location where the respective studies were conducted, and the baseline characteristics of participants between the case and control groups were comparable. When the OR and 95% CI values did not include the no-effect threshold of 1, a significant association with DED was indicated, and the factor was considered a risk factor of DED.

The synthesized data demonstrated that participant comorbidities (from the highest to the lowest strength of the association) of depression (OR, 1.60; 95% CI, 1.34 to 1.92), arthritis (OR, 1.56; 95% CI, 1.28 to 1.91), thyroid disease (OR, 1.49; 95% CI, 1.34 to 1.66), stroke (OR, 1.30; 95% CI, 1.21 to 1.39), and diabetes (OR, 1.15; 95% CI, 1.04 to 1.27) were risk factors of the DED; whereas hypertension (OR, 1.08; 95% CI, 0.99 to 1.18), smoking (OR, 1.14; 95% CI, 0.98, 1.33), and alcohol use (OR, 1.01; 95% CI, 0.84 to 1.20) were not associated with the DED (Table 2, Fig. 3).

Technical Validation

The gold-standard two-step process of two independent reviewers were conducted based on the consensus between the products of the two reviewers, and this was strictly implemented throughout the entire creation process of the DrDED. Two experienced reviewers with ophthalmologic expertise backgrounds screened the titles and abstracts, screened the full-texts, and independently acquired the data from eligible studies (Fig. 1). Each datapoint in the documents of the two reviewers was manually compared and validated using the statistical software, Review Manager (RevMan, Version 5.4, the Cochrane Collaboration, 2020). The accuracy of the extracted sample sizes, and the OR and 95% CI values were presented by generating a forest plot in both numeric and graphic forms. If any discrepancy was identified in the data and outliers reported by the two reviewers, the full text was carefully reviewed, and the data was corrected, accordingly. Any disagreements on the study eligibility was resolved through team discussion. Furthermore, an external methodologist conducted the final data validation by checking the study flow, logic and accuracy of all data.

Category	Data items
A. Study publication information	1. Study unique ID, including the subgroup labels where applicable
	2. The study record, which comprises of the last name of the first author and year of publication, and the full citation
	3. Full name of the first author
	4. Year of publication
	5. Name the journal where the study was published
B. Study characteristics	6. Type of study with the following options: cross-sectional, case-control, nest case-control, retrospective cohort, and prospective cohort studies.
	7. Risk of bias assessment
	8. Country where the study was conducted
	9. Sample population characteristics, such as population-based, hospital-based, and university-based population.
	10. Age of the study subjects
	11. Number of enrolled subjects
C. Definition and prevalence of the disease	12. Ethnicity
	13. Diagnosis standard used to define the DED in the study
	14. Prevalence of DED (point estimate and values of variability)
D. Risk factors	15. Central tendency estimate (in the form of OR) of each risk factor reported by the study*
	16. Lower and higher bounds of association between the risk factor and disease with 95% CIs for the ORs
	17. Subgroups, if any**.

Table 1. Structure of the DrDED. Note: DrDED, dataset on risk factors for dry eye disease; ID, identifier; DED, dry eye disease; OR, odds ratio; CI, confidence interval. *The data was reported as categorical variables, unless specified. **When the subgroup data is available, this will be entered in the DrDED in multiple rows and the study ID will be labelled with the number and information of the subgroups.

Patient factor	Study IDs	Overall assessment	Prospective cohort studies	Cross-sectional studies	Case-control studies	Retrospective cohort studies
Thyroid disease	3, 16, 29, 32, 48, 51, 58, 68, 85, 95, 97, 146	1.49-fold risk increase	1.27-fold risk increase	1.55-fold risk increase	—	No association
Arthritis	3, 10, 29, 32, 42, 48, 51, 68, 75, 81, 85, 95, 97, 103	1.56-fold risk increase	No association	1.51-fold risk increase	2.16-fold risk increase	—
Hypertension	22, 26, 32, 46, 48, 51, 57, 58, 67, 71, 75, 81, 85, 97, 103, 105, 127, 135, 147	No association	No association	No association	—	No association
Diabetes	3, 42, 51, 58, 67, 68, 71, 75, 81, 85, 92, 97, 103, 147	1.15-fold risk increase	No association	1.13-fold risk increase	1.41-fold risk increase	No association
Stroke disease	26, 75, 97	1.3-fold risk increase	—	1.3-fold risk increase	—	—
Smoking	3, 4, 18, 24, 28, 32, 36, 51, 54, 58, 60, 61, 71, 72, 81, 92, 97, 105, 147	No association	1.74-fold risk increase	No association	No association	No association
Alcohol use	24, 32, 57, 61, 81, 97	No association	—	No association	1.85-fold risk decrease (protective factor)	—
Depression	28, 32, 42, 56, 58, 75, 85, 92, 97, 98, 117, 135, 140, 146	1.6-fold risk increase	—	1.61-fold risk increase	1.45-fold risk increase	1.77-fold risk increase

Table 2. Overview for the subset data records in the DrDED (41 high-quality studies between 2000 and 2021 that reported one or more of the eight common risk factors for DED). The association and strength of the association between the reported risk factors and DED are listed in the table. For example: patients with thyroid disease were 1.49-fold more likely to develop DED, when compared to patients without the disease (OR, 1.49; 95% CI, 1.34–1.66); patients with diabetes were 1.15-fold more likely to develop DED (OR, 1.15; 95% CI, 1.04–1.27). DrDED, dataset on risk factors for dry eye disease; ID, identifier; --, no data; OR, odds ratio; CI, confidence interval.

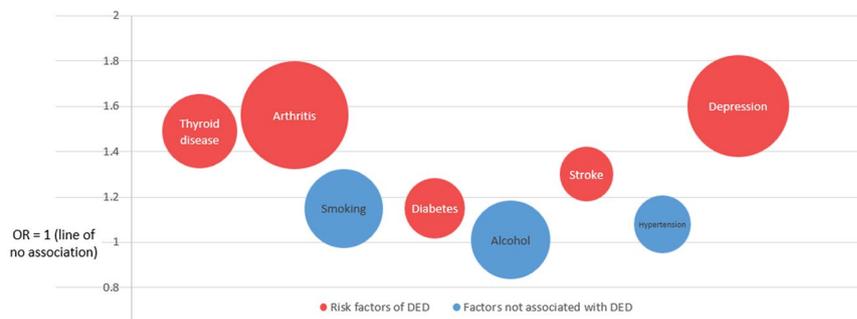


Fig. 3 Risk factors and the association strength for DED (a subset of high-quality studies between 2000 and 2020, which reported one or more of the eight common risk factors for DED). DED, dry eye disease. The further away the bubble was from the line of no association (OR [odds ratio] = 1), the stronger the association strength became. The red bubbles represent the risk factors, and the blue bubbles represent the factors not associated with DED.

Usage Notes

The DrDED can serve its users from two aspects. First, the data can be used to investigate the risk factors of DED through the subgroup meta-analysis of different age groups (younger vs. older), gender, countries/continents, sample population sources (population-based vs. specific institution-based), study sample size (large vs. small), the diagnosis criteria of DED, and other subgroup categories, when reported by primary studies. The user can use the data for further statistical analysis, such as Egger's test, in order to determine whether publication bias is present¹⁴. Second, the DrDED serves as a library and a research planning source for users. For clinical researchers who want to conduct their own epidemiological study to explore the risk factors of DED in a population, they can use the information in the dataset to assist in the study design, in terms of the study subject's inclusion criteria, determination of sample sizes, questionnaire design, and data to be collected. When needed, they can also refer to the details in the data analysis methods and background knowledge by reviewing the full texts of high-quality studies that were systematically searched and screened for the clinical research question.

The DrDED is a living dataset, because the information is updated through data search and acquisition on a yearly basis, and every time a well-known study is published. The procedures described in the present study will be applied, including the data validation, when adding data obtained from new publications to the DrDED, with focus on the risk factors for DED, thereby keeping the dataset up-to-date. The DrDED is publicly available on the journal platform in its electronic version. Its future up-to-date versions can be obtained by contacting the corresponding authors of the present study *via* E-mail.

Code availability

Not applicable.

Received: 10 June 2022; Accepted: 3 January 2023;

Published online: 11 January 2023

References

1. The definition and classification of dry eye disease. report of the Definition and Classification Subcommittee of the International Dry Eye WorkShop (2007). *Ocul Surf.* **5**, 75–92 (2007).
2. Shimazaki, J. Definition and Diagnostic Criteria of Dry Eye Disease: Historical Overview and Future Directions. *Invest Ophthalmol Vis Sci.* **59**, Des7–des12 (2018).
3. Gayton, J. L. Etiology, prevalence, and treatment of dry eye disease. *Clin Ophthalmol.* **3**, 405–412 (2009).
4. Chatterjee, S., Agrawal, D., Sanowar, G. & Kandoi, R. Prevalence of symptoms of dry eye disease in an urban Indian population. *Indian J Ophthalmol.* **69**, 1061–1066 (2021).
5. Tsubota, K. *et al.* Defining Dry Eye from a Clinical Perspective. *Int J Mol Sci.* **21** (2020).
6. Cai, Y., Wei, J., Zhou, J. & Zou, W. Prevalence and Incidence of Dry Eye Disease in Asia: A Systematic Review and Meta-Analysis. *Ophthalmic Res.* **65**, 647–658 (2022).
7. Vehof, J., Snieder, H., Jansonius, N. & Hammond, C. J. Prevalence and risk factors of dry eye in 79,866 participants of the population-based Lifelines cohort study in the Netherlands. *Ocul Surf.* **19**, 83–93 (2021).
8. Kuo, M. T., Fang, P. C., Kuo, S. F., Chen, A. & Huang, Y. T. Tear Proteomics Study of Dry Eye Disease: Which Eye Do You Adopt as the Representative Eye for the Study? *Int J Mol Sci.* **22** (2021).
9. Kuklinski, E. J. *et al.* Associations Between Systemic Omega-3 Fatty Acid Levels With Moderate-to-Severe Dry Eye Disease Signs and Symptoms at Baseline in the Dry Eye Assessment and Management Study. *Eye Contact Lens.* **47**, 2–7 (2021).
10. Shokr, H., Wolffsohn, J. S., Trave Huarte, S., Scarpello, E. & Gherghel, D. Dry eye disease is associated with retinal microvascular dysfunction and possible risk for cardiovascular disease. *Acta Ophthalmol.* **99**, e1236–e1242 (2021).
11. Rakofsky, J. J., Rakofsky, S. I. & Dunlop, B. W. Dry Those Crying Eyes: The Role of Depression and Antidepressants in Dry Eye Disease. *J Clin Psychopharmacol.* **41**, 295–303 (2021).
12. Tandon, R. *et al.* Association of dry eye disease and sun exposure in geographically diverse adult (≥ 40 years) populations of India: The SEED (sun exposure, environment and dry eye disease) study - Second report of the ICMR-EYE SEE study group. *Ocul Surf.* **18**, 718–730 (2020).
13. Craig, J. P. *et al.* TFOS DEWS II Definition and Classification Report. *Ocul Surf.* **15**, 276–283 (2017).
14. Craig, J. P. & Wang, M. T. Factors predisposing the Asian eye to dry eye disease. *Invest Ophthalmol Vis Sci.* **60**, 2746 (2019).

15. Wang, M. T. M., Muntz, A., Mamidi, B., Wolffsohn, J. S. & Craig, J. P. Modifiable lifestyle risk factors for dry eye disease. *Cont Lens Anterior Eye*. **44**, 101409 (2021).
16. Yang, *et al.* DrDED, a dataset on literature on risk factors for dry eye disease, *Figshare*., <https://doi.org/10.6084/m9.figshare.c.6009355.v1> (2022).
17. Higgins, J. P. T. & Altman, D. G. Chapter 8. Assessing risk of bias in included studies. In: Higgins J. P. T., Green S. (editors). *Cochrane Handbook for systematic Reviews of Interventions*. *Cochrane Collab*. **Version 5.1.0 [updated March 2011]**, (2011).

Acknowledgements

None.

Author contributions

W.Y. and Y.X. designed the study. All authors performed the database search, data collection, and study quality evaluation. W.Y. drafted the manuscript. All authors revised the manuscript, and approved the submission of the manuscript.

Funding

The study was supported by grants from the Youth Project of the Medical Scientific Research Foundation of Wuhan Municipal Health Commission (No. WX21Q33), the National Natural Science Foundation Youth Project (No. 81800802), the Science and Technology Innovation Program of Hunan Province (No. 2020SK50104), and the Scientific Research Fund project of Aier Eye Hospital Group (Nos. AR2210D2 and AF2004D10). The funders had no role in the design, data acquisition, or manuscript preparation of the study.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-01931-8>.

Correspondence and requests for materials should be addressed to Y.Y. or Y.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023