

# A BIBLIOMETRIC ANALYSIS OF THE RELATIONSHIP BETWEEN DIABETES AND ARTIFICIAL INTELLIGENCE\*

## DIYABET VE YAPAY ZEKÂ İLİŞKİSİ ÜZERİNE BİR BİBLİYOMETRİK ANALİZ

Denizhan DEMİRKOL<sup>1</sup> , Fatma Önay KOÇOĞLU<sup>2</sup> , Şamil AKTAŞ<sup>3</sup> , Çiğdem EROL<sup>4,5</sup> 

<sup>1</sup>İstanbul University, Institute of Science, Department of Informatics, İstanbul, Türkiye

<sup>2</sup>Muğla Sıtkı Koçman University, Faculty of Engineering, Department of Software Engineering, Muğla, Türkiye

<sup>3</sup>İstanbul University, İstanbul Faculty of Medicine, Department of Underwater and Hyperbaric Medicine, İstanbul, Türkiye

<sup>4</sup>İstanbul University, Department of Informatics, İstanbul, Türkiye

<sup>5</sup>İstanbul University, Science Faculty, Department of Biology, İstanbul, Türkiye

**ORCID IDs of the authors:** D.D. 0000-0003-1215-0404; F.O.K. 0000-0002-1096-9865; Ş.A. 0000-0002-9242-3179; Ç.E. 0000-0002-5057-7145

**Cite this article as:** Demirkol D, Kocoglu FO, Aktas S, Erol C. A bibliometric analysis of the relationship between diabetes and artificial intelligence. J Ist Faculty Med 2022;85(2):249-57. doi: 10.26650/IUITFD.928111

### ABSTRACT

**Objective:** The applications of artificial intelligence in the field of medicine are progressing at a remarkable speed. Within the scope of this study, it is aimed to make a detailed examination of the studies published on diabetes and artificial intelligence, to determine any trends in these studies over time, to examine which subjects have been researched more and to explain the global interest in the subject.

**Material and Methods:** In this study, 2534 studies published between 1985 and 2020 in the field of diabetes and artificial intelligence were analyzed using the R programming language through bibliometric data obtained from the Scopus database. Correlation analysis, ANOVA and regression analysis were performed to determine the relationship between the number of articles and years.

**Results:** According to the analysis results, the number of publications between 1985 and 2015 was 604 and over the last 5 years, the number of publications have tripled to 1930. It was found that the country with the highest number of publications with 358 publications and 10426 citations was the United States of America (USA). Moreover, in the analyzed studies, the most used keywords and the use of these words together was also examined and the top 10 source platforms where the studies were published most have been presented in the study. According to regression analysis, it can be predicted that the number of articles to be published for 2021 is 242.

**Conclusion:** As a result of the analysis in this study, it was determined that artificial intelligence and diabetes applications have

### ÖZET

**Amaç:** Yapay zekânın sağlık alanındaki uygulamaları dikkat çekici hızla ilerlemektedir. Bu çalışma kapsamında diyabet ve yapay zekâ konusunda yayınlanan çalışmaların detaylı bir incelemesinin yapılması, çalışmaların zaman içindeki eğiliminin belirlenmesi, hangi konularda daha fazla araştırma yapıldığının incelenmesi ve konuya ilgili küresel ilginin açığa çıkarılması amaçlanmıştır.

**Gereç ve Yöntem:** Bu çalışmada 1985-2020 yılları arasında diyabet ve yapay zekâ alanında yayınlanan 2534 çalışma Scopus veri tabanından elde edilen bibliyometrik verilerle R programlama dili kullanılarak analiz edilmiştir. Yıllara göre makale sayısı arasındaki ilişkiyi açığa çıkarmak için, korelasyon analizi, varyans analizi (ANOVA) ve regresyon analizi gerçekleştirilmiştir.

**Bulgular:** Yapılan analiz sonuçlarına göre, 1985-2015 yılları arasında üretilen yayın sayısının 604 olduğu ve son 5 yıldaki yayın sayısının bu sayıyı üç katlayarak 1930'e çıkardığı tespit edilmiştir. Bununla birlikte, 358 yayıyla en çok yayın yapan ve 10426 atıfla en çok atıf alan ülkenin, Amerika Birleşik Devletleri olduğu saptanmıştır. Ayrıca, incelenen araştırmalarda en çok kullanılan anahtar kelimeler ve bu kelimelerin bir arada kullanımı ve çalışmaların en çok yayınlandığı ilk 10 kaynak platform da araştırmada sunulmuştur. Gerçekleştirilen regresyon analizine göre 2021 yılı için yayınlanabilecek makale sayısının 242 olarak tahmin edilmiştir.

**Sonuç:** Sonuç olarak, yapay zekânın altın çağını yaşadığı günümüzde, yapay zekâ ve diyabet uygulamalarının, küresel bazda önemli araştırma konularından biri hâline geldiği saptanmış olup, diyabetti önlemek veya erken dönemde teşhis etmek için

\* This study was presented orally at the 7<sup>th</sup> International Management Information Systems Conference.

Corresponding author/Iletişim kurulacak yazar: denizhan.demirkol.2018@ogr.edu.tr

Submitted/Başvuru: 26.04.2021 • Revision Requested/Revizyon Talebi: 30.07.2021 •

Last Revision Received/Son Revizyon: 08.01.2022 • Accepted/Kabul: 13.01.2022 • Published Online/Online Yayın: 08.02.2022



Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

become one of the important global research topics in today's golden age of artificial intelligence, and it was also determined that there is an urgent need for artificial intelligence supported scientific studies to prevent diabetes or diagnose diabetes in the early period.

**Keywords:** Diabetes mellitus, artificial intelligence, bibliometric analysis

## INTRODUCTION

The incidence of diabetes is increasing day by day with the spread of obesity and sedentary lifestyles and it has become a global health problem. In fact, it could be said that diabetes is the health crisis of the century (1). According to the International Diabetes Federation's 9<sup>th</sup> Diabetes Atlas data in 2019, since its first report in 2000, the estimated number of patients with diabetes (type 1 and type 2 combined, both diagnosed and undiagnosed) was 151 million people aged between 20 to 79 years (4.6% of the global population) and increased to 463 million people (9.3% of the global population). The report predicts that this number, which was estimated to be 463 million in 2019, will reach 578 million in 2030, and if sufficient action plans are not implemented for this pandemic, this number will reach 700 million (10.9%) in 2045. In addition, the number of deaths from diabetes and its complications was estimated to be 4.2 million in 2019. In the same report, it was stated that 10 percent of global health expenses are spent on diabetes, while 3 out of every 4 people with diabetes (79%) live in low-middle income countries (2). The prevalence of diabetes is also increasing in Turkiye, parallel to its global increase. The first of the studies conducted to determine the prevalence of diabetes in Turkiye is TURDEP I (Turkiye Diabetes, Hypertension, Obesity and Endocrinology Diseases Prevalence Study-I) study (3). According to the results of the study, the diabetes prevalence was found to be 7.2% (24788 subjects, age $\geq$ 20, mean  $40.88\pm14.86$ ). Moreover, according to the results of the TEKHARF 2013 (Turkish Adult Risk Factor survey 2013) study, which was conducted to explain the change in the prevalence of diabetes across the country in the last 12 years (768 subjects, age range ~40-80), it was found that the overall prevalence of diabetes increased by 80%, which corresponds to an annual increase of 5% (4). According to the Turkiye Chronic Kidney Disease Prevalence Research (the Chronic Renal Disease in Turkiye - CREDIT) results (10872 participants, mean age  $40.5\pm16.3$  years) diabetes prevalence was 14.3% in women, and 10.9% in men and it is overall reported as 12.7% (5). According to the TURDEP-II study, the prevalence of diabetes in adults 20 years and older was determined as 16.5% and it was stated that according to the data of 2010, there were 6.5 million diabetic people in Turkiye (6).

Long-term complications of diabetes include retinopathy, potentially leading to impaired vision; nephropathy,

yapay zekâ destekli bilimsel çalışmalarla yoğun bir ihtiyacın bulunduğu tespit edilmiştir.

**Anahtar Kelimeler:** Diyabet, yapay zekâ, bibliyometrik analiz

leading to kidney failure; peripheral neuropathy with risk of foot ulcer, Charcot joints; amputation and autonomic neuropathy causing gastrointestinal, genitourinary and cardiovascular symptoms and sexual dysfunction. Patients with diabetes have an increased incidence of atherosclerotic cardiovascular, peripheral arterial and cerebrovascular diseases. However, if proper diabetes management is provided, these serious complications can be delayed or avoided altogether (2). Research results demonstrate that, with the increasing incidence of diabetes, early detection and prevention of both diabetes and diabetes complications are critical.

Artificial intelligence (AI) has become one of the most important fields of study worldwide. AI is defined as the ability of a computer or a computer-controlled machine to perform tasks related to higher mental processes such as reasoning, inferring and learning from past experiences, which are generally assumed to be human-specific qualities (7). AI applications have become prominent especially in the field of medicine. It is one of the important research areas that will benefit humanity the most and research is progressing at a remarkable pace (8). To illustrate this, with the aging population in the world, chronic diseases such as diabetes, cancer, and dementia have tended to increase and physicians, nurses and all employees who are in direct contact with the patient, working in the health sector, make a serious effort to adapt to the patient's condition when they come across such patients. It is obvious that in such cases, AI will be very useful as a decision supporting system (9). Digital medicine and related research have the potential to continuously monitor patients' symptoms, physiological data, behavior, and lifestyle remotely through AI, wearable devices, sensors and smartphone technologies. Thus, it has revolutionary effects in the prevention of diabetes, from the creation of personal nutrition plans to the regulation of the necessary diabetes treatment (10).

It has been stated that AI can help reduce the global prevalence of diabetes of 8.8% by changing the way diabetes is prevented, detected, and managed. Furthermore, it was stated that it would positively affect the decision-making process and thus cause a paradigm change in the management of diabetes based on data (11). AI technologies with machine learning algorithms have started to be used for the prediction and management of diabetes and related syndromes from the diagnosis of

diabetes, diabetic retinopathy and diabetic nephropathy to blood glucose control, and studies in this area are increasing day by day (12-16).

Studies on the use of AI in the diagnosis and management of diabetes have become increasingly common in the literature (17). The development and progress of this increasing number of studies, the level of interest around the world, and especially the subjects of these research studies, have been an object of curiosity. Regardless of the field, before conducting a new scientific study, it is necessary to determine which new trends and gaps are in the field, which methods and models are used, the aspects of the study to be conducted that differ from other studies and what innovations it can bring to the literature. Literature studies are carried out before the application and it is stated that it is possible to carry out these studies more technically using bibliometric data (18).

Within the scope of this study, it is aimed to make a detailed examination of the studies published on diabetes and AI, to determine the trend of the studies over time, as stated above, to examine which subjects have been researched more and to explain the global interest in the subject. Thus, it is aimed to present a preliminary literature that will guide researchers who want to conduct similar studies in the literature. Below, the data and the method used in the study, the findings obtained, and the inferences about the past, present and future of diabetes studies related to AI based on the findings obtained are shared respectively.

## MATERIALS AND METHODS

In line with the purpose of the study, the bibliometric analysis technique was used. Bibliometry is defined as being a part of research evaluation methodology and the numerical analysis of the publications produced by individuals or institutions in a specific area in a specific region and the relations between these publications (19). This type of analysis is based on the identification of publications in a particular subject area (20). Within the scope of bibliometric analysis, processes such as quantitative evaluation of publications and citation data in a field of science, measuring the progress in the related field, determining conceptual relations and related network maps, and interpreting trends are carried out. In more general terms, bibliometric analysis can be examined under two main headings - performance analysis and mapping. Performance analysis is based on numerical calculations such as the number of publications and citations by country and author, and determination of the most frequently used words. Mapping aims to reveal the relationship between the mentioned variables. Various measures such as the number of publications and citations per year, citation distribution, number of citations per publication, number of authors per publication, collaboration index,

and h-index are used for performance analysis. Citation analysis, co-citation analysis, bibliographic coupling, co-word analysis, and co-authorship analysis are used for mapping (21).

In order to perform bibliometric analysis, bibliometric data must first be obtained. There are many databases such as Web of Science, Scopus, PubMed, Google Scholar, Embase, and SpringerLink, all of which can be used to obtain data. Scopus is a database of studies in high-quality scientific journals, conference proceedings and international books. It ensures that only the highest quality data is indexed through rigorous content selection and re-evaluation by an independent Content Selection and Advisory Board. Following the same rigorous standards, it also creates a comprehensive author and institution profile with the support of profiling algorithms (22). For these reasons, the Scopus database was chosen to obtain the necessary data for this study.

Within the scope of the study, the bibliometric data of academic studies in accordance with the following constraints were obtained from this database.

- Including "diabetes-artificial intelligence" or "diabetes-machine learning" word groups in the title, summary, or keywords
- Published in the form of article, book series, books and conference proceedings
- Published/accepted for publication between 1985-2020
- Published in English

Regarding 2534 studies in total:

- Quote information
  - Author (s) name (s), author (s) ID (s), study title, year, reference title, volume-issue-page information, citation number, document type, stage in the publication process, access type, DOI number.
- Bibliographic information
  - Institution, ISSN, PubMed ID, publisher, editor, original language, responsible author address information, abbreviated source title.
- Summary and keyword information
  - Summary, keywords, indexed keywords.
- Supporting information
  - Number, acronym, sponsor, financing text.
- Other information
  - Trade name and manufacturers, accession numbers, conference name, attached references.

Although the same database was searched with different word groups, some studies were repeated as machine learning is a subfield of AI. A total of 668 duplicate studies (determined according to DOI numbers), were excluded from the data set used in the analysis. The study carried

out was limited to the database used and the preferred word groups.

The term AI includes many sub-branches such as machine learning, data mining, evolutionary algorithm, expert systems, etc. Therefore, while obtaining bibliometric data on the articles, different combinations of the word groups related to these sub-branches could be used together with AI. In order to obtain data in this study, the researchers limited the search word groups to AI-diabetes and diabetes-machine learning because all the branches mentioned are sub-branches of AI as stated. AI was chosen in this study since it is a very inclusive concept. While the sections in which the word groups to be searched in the articles were limited to only the title or keywords, the abstract section is considered as the article section where the word groups are searched in order to obtain the data of the studies in the sub-branches, and the search space was expanded. At this point, it was thought that the concept of AI would be included at least in the summary section for studies on expert systems, data mining, etc. Another situation is minimizing the repetition of the articles the data of which will be extracted. The data of the articles listed in the search with each word group is a subgroup dataset. The data set, which is the result of combining all these sub-datasets, is used for the analysis. Therefore, if each sub-branch word group is used to obtain data, some articles will repeat in the data collected for each word group, since the aforementioned fields are related to each other. This repetition is minimized with the selected inclusive word groups. Machine learning is also among the mentioned sub-fields, but the diabetes-machine learning word group was also used in this study. At this point, the researchers used this word group especially in order to see which algorithms would come to the fore in the articles on the development of the prediction model.

Codes were written in R programming language by using RStudio editor for bibliometric data analysis (23,24). With the R programming language, code packages prepared for bibliometric analysis studies, as well as for many different analysis types, have been developed and various functions are included in these packages. The packages and functions used in this study are given below:

#### Package:

- **bibliometrix** (25): The bibliometrix package is a tool that contains many functions that can be used to perform quantitative research in bibliometrics and scientometrics. The functions in this package generally provide solutions for three purposes. Making bibliometric data suitable for analysis with R programming language, performing bibliometric analysis of a publication data set, creating matrices needed for co-citation, coupling, collaboration and common word analysis.

- **tidyverse** (26): The tidyverse package is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures.

Functions used from bibliometrix package:

- **readFiles()**: This function is used to assign the exported SCOPUS files into a large character object.
- **convert2df()**: This function is used to import and convert bibliographic files and API objects.
- **biblioAnalysis()**: This function is used to perform bibliometric analysis.
- **biblioNetwork()**: This function has been used to create bibliographic networks.
- **networkPlot()**: This function was used to plot bibliographic networks.

Functions used from the tidyverse package:

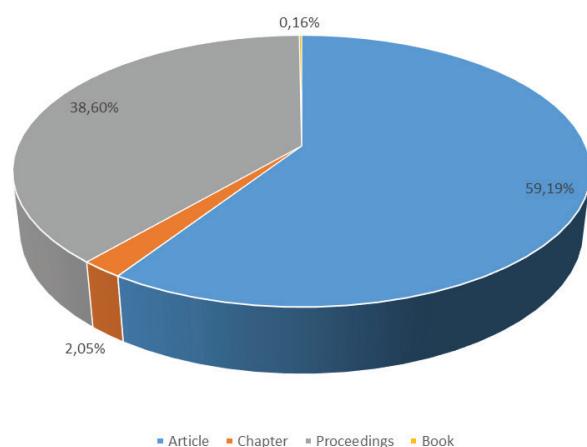
- **duplicated()**: The function was used to identify repetitive studies in the data set.

In addition, correlation analysis, one-way analysis of variance (ANOVA) and regression analysis were carried out in order to determine the relationship between the number of articles by years. All the statistical analyses were performed using SPSS 21.0 software. A value with a p-value of  $\leq 0.05$  were considered statistically significant.

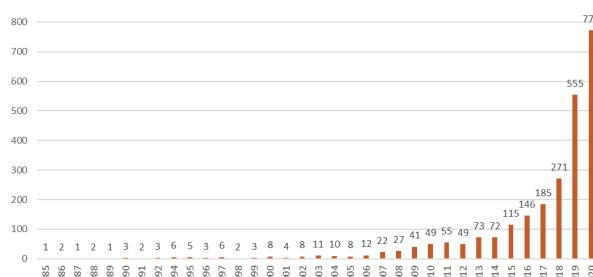
## RESULTS

The findings obtained as a result of the bibliometric analysis study performed within the steps specified in the method section have been shared under this title. 1500 of the 2534 publications included in the analysis are articles, 978 are conference papers, 52 are book chapters, and 4 are books. The percentage distribution of the publications is given in Figure 1.

The frequency distribution of the publications included in the analysis by years was also obtained. While the num-



**Figure 1:** Percentage distribution by type of publications



**Figure 2:** Counts depending on the years

\*SCP: Single Country Publication, MCP: Multi Country Publication

to the total number of multi-author articles (27, 28). The cooperation index, calculated accordingly, is 4.21.

Findings about the relationship between the country and the number of publications are given in Table 2 and Figure 3. Findings on the top 10 producing countries are also shared in the table. Accordingly, the country producing the most publications with 358 publications was the United States of America (USA), while the other countries in the top 10 were China, India, United Kingdom, Korea, Spain, Italy, Australia, Germany, and France. When the multi-country collaborative publication rates (MCP\_Ratio)

**Table 1:** Author-publication number relationship

Total number of authors	Number of publications with a single author	Number of publications per author	Number of authors per publication	Collaboration index
10,274	109	0.247	4.05	4.21

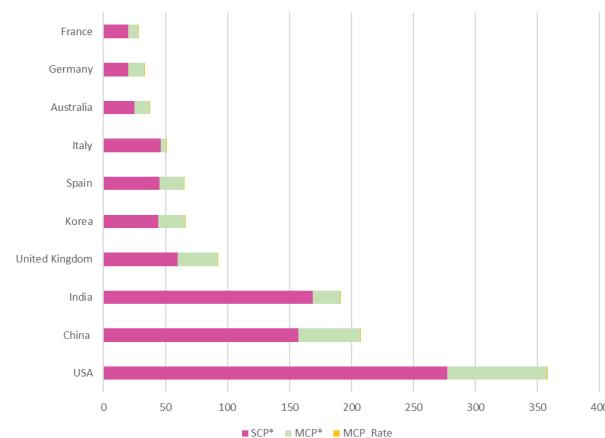
**Table 2:** Country-publication relationship

Country	Number of publications	SCP	MCP	MCP_Rate
USA	358	277	81	0.226
China	207	157	50	0.242
India	191	169	22	0.115
United Kingdom	92	60	32	0.348
Korea	66	44	22	0.333
Spain	65	45	20	0.308
Italy	51	46	5	0.098
Australia	37	25	12	0.324
Germany	33	20	13	0.394
France	28	20	8	0.286

SCP: Single country publication, MCP: Multi country publication

ber of publications between 1985 and 2015 was 604, the number of publications in the last 5 years has tripled to 1930. Publication counts by years can be seen in Figure 2. Although the annual average increase rate was calculated as 20.93%, when Figure 2 is analyzed, it is seen that the number of publications has increased especially in the last 5 years and reached the highest level with the publication of a total of 773 studies in 2020.

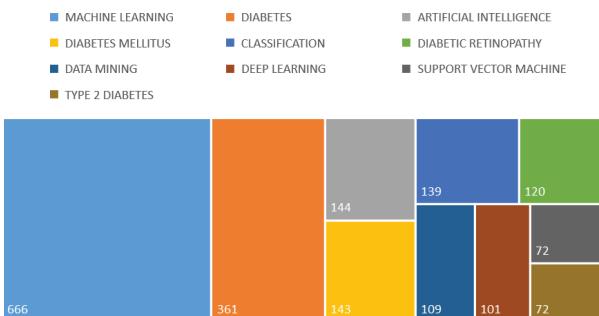
Findings regarding the relationship between author and publication number are given in Table 1. Accordingly, the total number of authors was 10,274, while the number of publications with a single author was 109. Although the majority of the studies had multiple authors, the number of authors per article was 4.05. The Collaboration Index (CI) is the ratio of the total authors of multi-author articles



**Figure 3:** Country-publication relationship chart

**Table 3:** Number of citations within the scope of the country-publication relationship

Country	Total number of citations	Number of citations per publication
USA	10426	29.12
China	2669	12.89
India	2330	12.20
United Kingdom	1639	17.82
Turkiye	1352	54.08
Australia	1346	36.38
Singapore	1137	54.14
Korea	1118	16.94
Israel	1077	59.83
Italy	1002	19.65



**Figure 4:** Frequency distribution of keywords

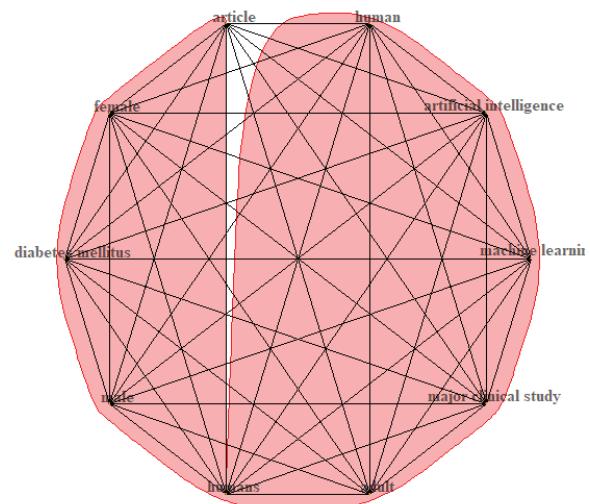
**Table 4:** Top 10 sources for publications

Platform	Number of publications
Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	116
Journal of Diabetes Science and Technology	55
Studies in Health Technology and Informatics	46
Advances in Intelligent Systems and Computing	45
Artificial Intelligence in Medicine	43
Communications in Computer and Information Science	36
ACM International Conference Proceeding Series	34
Plos One	33
Journal of Biomedical Informatics	31
Procedia Computer Science	27

among all the publications of the countries was examined, Germany, United Kingdom, Korea, Australia, and Spain were in the top 5.

The findings regarding the number of citations obtained depending on the country-publication relationship are given in Table 3. Accordingly, the USA was the country with the highest number of citations with 10426 citations. When the table is examined, the top 10 countries with the most citations were the USA, China, India, United Kingdom, Turkiye, Australia, Singapore, Korea, Israel, and Italy. Although the USA publications had the most citations, when the number of citations per publication was examined, Israel came first with a rate of 59.83%.

In the analyzed studies, the most used keywords and the use of these words together were also examined. Accord-



**Figure 5:** Relationship chart between keywords

**Table 5:** Summary of the created regression model

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	Std. Error of the estimate	Change statistics			
					R <sup>2</sup> change	F change	df1	df2
Model 1	0.613 <sup>a</sup>	0.376	0.357	127.917	0.376	20.474	1	34

a: Dependent variable: articles, df: degrees of freedom

**Table 6:** ANOVA<sup>a</sup> results

Model	Sum of squares	df	Mean square	F	p
Model 1	Regression	1	335019.287	20.474	0.000 <sup>b</sup>
	Residual	34	16362.802		
	Total	35			

a: Dependent variable: articles, b: Predictors: (constant), years, df: degrees of freedom

**Table 7:** Results regarding the prediction of the number of articles by years

Model	Unstandardized coefficients		Beta	95.0% Confidence interval for B		
	B	Std. error		t	p	Lower bound
Model 1	(Constant)	-101.406	43.543	-2.329	0.026	-189.897
	Years	9.286	2.052			

t: test value, p: probability value, B: beta coefficient

ingly, the most used keywords were "Machine Learning", "Diabetes", "Artificial Intelligence", "Diabetes Mellitus", "Classification", "Diabetic Retinopathy", "Data Mining", "Deep Learning", "Support Vector Machine", and "Type 2 Diabetes". The frequency distribution of keywords is given in Figure 4, and the situation of using words together is given in Figure 5.

The top 10 source platforms where the studies were published most are given in Table 4.

The results of the regression model created according to the number of articles by years are given in Table 5. The correlation between the number of articles by years was positive, linear and significant ( $r=0.613$ ,  $p\leq 0.01$ ).

The output of the ANOVA, which tests the significance of the regression model, is given in Table 6. When the ANOVA results regarding the prediction of the number of articles by years were evaluated, there was significance ( $F(1,34)=20.474$ ,  $p\leq 0.01$ ). The regression model established for the prediction of the number of articles by years was significant.

The results of the prediction of the number of articles for the year 2021 are given in Table 7.

When the coefficients of the model were evaluated, 37.6% of the total variance was explained and the years

had a significant effect on the number of articles ( $\beta=9.286$ ;  $\beta_{std}=0.613$ ;  $t(34)=4.525$ ;  $p:0.000\leq 0.001$ ).

The regression equation obtained for the model was:

$$Y = -101.406 + 9.286X.$$

The calculation to predict the number of articles for the year of 2021 (the 37th year) was:

$$Y = -101.406 + 9.286(37) = -101.406 + 343.582 \approx 242.$$

Accordingly, it can be predicted that the number of articles to be published for 2021 is 242.

## CONCLUSION

AI technologies continue to transform our habits and the health sector is taking its share from this transformation. The importance of digital transformation is increasing with the integration of AI technologies in the prevention, monitoring and management of diabetes, which is one of the most important public health problems of today. With the effect of many factors such as obesity, unhealthy diet and sedentary lifestyle, diabetes has become a pandemic and has become a global public health problem. Therefore, it can be said that there is an urgent need for AI supported scientific studies to prevent diabetes or diagnose it at an early stage.

In this study, a bibliometric analysis study was carried out with the data of 1029 studies, which included "diabetes-artificial intelligence" or "diabetes-machine learning" word groups in the title, a summary or keywords, and which were published or accepted for publication between 1985-2020 in the form of articles, book chapters and conference papers. As a result of the analysis, quite remarkable findings were obtained. It was determined that the number of studies conducted since 2015 was triple the number of studies published between 1985 and 2015. These findings are not surprising given that AI has had its second golden age since 2010 (29). Moreover, when the number of studies carried out up until the end of 2020 is examined, it can be stated that this increased interest will continue in the 2021 according to regression analysis.

When the most used keywords in the studies and the situations of using these words together were examined, it was seen that the term "machine learning" was used more frequently than AI. Accordingly, it can be said that instead of smart applications such as wearable technologies and remote monitoring systems, the studies were mainly based on predictive analysis based on the data set (in other words, machine learning from experience). Likewise, the concept of classification was seen in keywords. This also supports the idea that the studies were prediction oriented. Although there are three basic methods in data mining - classification, clustering and association analysis - the classification method was at the forefront of the study results. In this respect, it can be thought that studies primarily focus on the benefit of prevention rather than monitoring of the disease. A more technical result is the keyword obtained especially for the algorithms used in machine learning. These are artificial neural network based deep learning algorithms and they support vector machines. Although there are many algorithms used for classification, the prominence of deep learning and support vector machines algorithms was interpreted as these two algorithms are used extensively in studies in this field. Although deep learning is the algorithm that has recently gained interest and is being used widely, it is surprising that support vector machines are among the keywords. It can be said that this algorithm is frequently preferred due to the nature of the data set used in studies in the relevant field and the desired results.

When the number of publications was examined, it was seen that the country that produced the most publications in this field with 358 publications and the most cited country with 10426 citations was the USA. The fact that the USA is one of the most preferred countries in scientific publication production, projects and collaborations and attaches importance to multidisciplinary work shows that this result is not surprising. However, when the num-

ber of citations per publication was examined, it was determined that the USA lags behind Israel (29.12 vs 59.83 respectively). Additionally, Singapore and Turkiye have a high number of citations per publication. However, if we consider that there are few publications in these countries, it may have been highly cited by other authors because of the lack of options.

It was determined that the published publications were predominantly multi-author and the average number of people per publication was above four. From this point of view, it was seen that teamwork especially stands out in this area. Considering the subject in general, it can be stated that the partnership of teams working in the field of health and informatics has gained importance in such research.

According to the study, Turkiye is unfortunately not among the top 10 countries producing the most publications in this field. In Turkiye, whose success in the field of health has been proven worldwide (30), the importance of health tourism is highlighted and where there is a desire to make a breakthrough with various investments (31), it needs studies to be carried out in this field for Turkiye to take its place among the leading countries. Despite the shortage of publications, it seems that Turkiye is among the top 10 countries in terms of the number of citations received. There could be many reasons why publications are cited more frequently. These reasons should be investigated in further studies separately.

When it comes to the limitations of this study, the study was carried out adhering to the data set obtained from a single database. For this reason, different databases may be considered in future studies. While listing the publications, a search was made for diabetes, machine learning and AI word groups. In future studies, these words could be diversified, customized or different combinations could be tried and the work could be expanded by accessing the data of different publications.

The authors declare that there is no conflict of interest and no financial support was received for the conduction of this research.

---

**Peer Review:** Externally peer-reviewed.

**Author Contributions:** Conception/Design of Study- D.D., F.Ö.K., §.A., Ç.E.; Data Acquisition- D.D., F.Ö.K.; Data Analysis/ Interpretation- D.D., F.Ö.K., §.A., Ç.E.; Drafting Manuscript- D.D., F.Ö.K., §.A., Ç.E.; Critical Revision of Manuscript- D.D., F.Ö.K., §.A., Ç.E.; Approval and Accountability- D.D., F.Ö.K., §.A., Ç.E.

**Conflict of Interest:** Authors declared no conflict of interest

**Financial Disclosure:** This study was partially funded by Scientific Research Projects Coordination Unit of İstanbul University. Project number: 37697.

## REFERENCES

1. Coşansu G. 21. Yüzyılın Sağlık Krizi: Diyabet. Florence Nightingale J Nurs 2014;17(2):115-22.
2. International Diabetes Federation. IDF Diabetes Atlas. Vol. 9th Edition. 2019 [cited 2021 May 11]. Available from: <https://www.diabetesatlas.org/en/resources/>
3. Satman I, Yilmaz T, Sengül A, Salman S, Salman F, Uygur S, et al. Population-based study of diabetes and risk characteristics in Turkey: Results of the Turkish Diabetes Epidemiology Study (TURDEP). Diabetes Care 2002;25(9):1551-6. [\[CrossRef\]](#)
4. Onat A, Çakır H, Karadeniz Y, Dönmez İ, Karagöz A, Yüksel M, et al. Turkish Adult Risk Factor survey 2013: rapid rise in the prevalence of diabetes. Turk Kardiyol Dernegi Arsivi-Arch Turk Soc Cardiol 2014;42(6):511-6. [\[CrossRef\]](#)
5. Süleymanlar G, Utaş C, Arınsoy T, Ateş K, Altun B, Altıparmak MR, et al. A population-based survey of Chronic REnal Disease In Turkey-the CREDIT study. Nephrol Dial Transplant 2011;26(6):1862-71. [\[CrossRef\]](#)
6. Satman I, Omer B, Tutuncu Y, Kalaca S, Gedik S, Dinccag N, et al. Twelve-year trends in the prevalence and risk factors of diabetes and prediabetes in Turkish adults. Eur J Epidemiol 2013;28(2):169-80. [\[CrossRef\]](#)
7. Nabiiev N. Yapay zeka. 4th ed. Ankara; 2012.
8. Buch V, Varughese G, Maruthappu M. Artificial intelligence in diabetes care. Diabet Med 2018;35(4):495-7. [\[CrossRef\]](#)
9. Bulut M. Önsöz. In: Bulut M, Dilmen N, Bora E, Gezer M, Erol Ç, Türker Şener L, editors. Sağlıklı Yapay Zekâ. İstanbul, Turkey: Çağlayan Kitabevi; 2019.
10. Fagherazzi G, Ravaud P. Digital diabetes: Perspectives for diabetes prevention, management and research. Diabetes Metab 2019;45(4):322-9. [\[CrossRef\]](#)
11. Ellahham S. Artificial Intelligence: The Future for Diabetes Care. Am J Med 2020;133(8):895-900. [\[CrossRef\]](#)
12. Huang Y, McCullagh P, Black N, Harper R. Feature selection and classification model construction on type 2 diabetic patients' data. Artif Intell Med 2007;41(3):251-62. [\[CrossRef\]](#)
13. Makino M, Yoshimoto R, Ono M, Itoko T, Katsuki T, Koseki A, et al. Artificial intelligence predicts the progression of diabetic kidney disease using big data machine learning. Sci Rep 2019;9(11862):1-9. [\[CrossRef\]](#)
14. Özmen EP, Özcan T. Diagnosis of diabetes mellitus using artificial neural network and classification and regression tree optimized with genetic algorithm. J Forecast 2020;39(4):661-70. [\[CrossRef\]](#)
15. Wong TY, Sabanayagam C. Strategies to tackle the global burden of diabetic retinopathy: from epidemiology to artificial intelligence. Ophthalmologica 2020;243(1):9-20. [\[CrossRef\]](#)
16. Zhao M, Jiang Y. Great expectations and challenges of artificial intelligence in the screening of diabetic retinopathy. Eye 2020;34(3):418-9. [\[CrossRef\]](#)
17. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: literature review. J Med Internet Res 2018;20(5):e10775. [\[CrossRef\]](#)
18. Koçoğlu F. "Endüstri 4.0" Konusu Üzerine R Programlama Dili ile Bibliyometrik Analiz. In: Başal, Handan Asude, Ulutürk Y, Öner NK, editors. Modern dönemde edebiyat, eğitim, iktisat ve mühendislik. Ankara: Berikan Yayınevi; 2018. p. 859-89.
19. TÜBİTAK ULAKBİM, Cahit Arf Bilgi Merkezi. Bibliyometrik Analiz Sıkça Sorulan Sorular. [cited 2020 Aug 27]. Available from: <https://cabim.ulakbim.gov.tr/bibliyometrik-analiz/bibliyometrik-analiz-sikca-sorulan-sorular/>
20. Ellegaard O, Wallin JA. The bibliometric analysis of scholarly production: How great is the impact? Scientometrics 2015;105(3):1809-31. [\[CrossRef\]](#)
21. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and guidelines. J Bus Res 2021;133:285-96. [\[CrossRef\]](#)
22. Baas J, Schotten M, Plume A, Cote G, Karimi R. Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. Quant Sci Stud 2020;1(1):377-86. [\[CrossRef\]](#)
23. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018. Available from: <https://www.R-project.org/>
24. RStudio Team. RStudio: Integrated Development Environment for R Boston, MA: RStudio, PBC; 2020. Available from: <http://www.rstudio.com/>
25. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. J Informetr 2017;11(4):959-75. [\[CrossRef\]](#)
26. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the Tidyverse. J Open Source Softw 2019;4(43):1686. [\[CrossRef\]](#)
27. Elango B, Rajendran P. Authorship trends and collaboration pattern in the marine sciences literature : a scientometric study. Int J Inf Dissem Technol 2012;2(3):166-9.
28. Koseoglu MA. Mapping the institutional collaboration network of strategic management research: 1980-2014. Scientometrics 2016;109(1):203-26. [\[CrossRef\]](#)
29. Kaynak O. The golden age of Artificial Intelligence. Discov Artif Intell 2021;1(1). [\[CrossRef\]](#)
30. Agartan TI. Politics of success stories in the path towards Universal Health Coverage: The case of Turkey. Dev Policy Rev 2021;39(2):283-302. [\[CrossRef\]](#)
31. Buse C, Unluonen, K. Economic evaluation of health tourism in Turkey. Journal of Tourismology 2020;6(1):99-109. [\[CrossRef\]](#)