Heliyon 9 (2023) e20524

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Systematic review and meta-analysis

A systematic review and future research agenda on detection of polycystic ovary syndrome (PCOS) with computer-aided techniques

Sayma Alam Suha [a], Muhammad Nazrul Islam [b,*]

[a] Department of Computer Science and Engineering, Bangladesh Army International University of Science and Technology, Cumilla, Bangladesh
[b] Department of Computer Science and Engineering, Military Institute of Science and Technology, Dhaka, Bangladesh

ARTICLE INFO

Keywords:
Polycystic ovary syndrome (PCOS)
Computer-assisted methods
Systematic literature review (SLR)
Data synthesis
Future research scopes

ABSTRACT

Polycystic Ovary Syndrome (PCOS) is among the most prevalent endocrinological abnormalities seen in reproductive female bodies posing serious health hazards. The correctness of interpreting this condition depends heavily on the wide spectrum of associated symptoms and the doctor's expertise, making real-time clinical detection quite challenging. Thus, investigations on computer-aided PCOS detection systems have recently been explored by several researchers worldwide as a potential replacement for manual assessment. This review study's objective is to analyze the relevant research works on computer-assisted methods for automatically identifying PCOS through a systematic literature review (SLR) methodology as well as investigate the research limitations and explore potential future research scopes in this domain. 28 articles have been selected using the PRISMA approach based on a set of inclusion-exclusion criteria for conducting the review. The data synthesis of the selected articles has been conducted using six data exploration themes. As outcomes, the SLR explored the topical association between the studies; their research profiles; objectives; data size, type, and sources; methodologies applied for the detection of PCOS; and lastly the research outcomes along with their evaluation measures and performances. The study also highlights areas for future research directions examining the study gaps to enhance the current efforts for autonomous PCOS identification; such as integrating advanced techniques with the current methods; developing interactive software systems; exploring deep learning and unsupervised machine learning techniques; enhancing datasets and country context; and investigating more unknown factors behind PCOS. Thus, this SLR provides a state-of-the-art paradigm of autonomous PCOS detection which will support significantly efficient clinical assessment, diagnosis and treatment of PCOS.

1. Introduction

Polycystic Ovary Syndrome (PCOS) which affects almost 15% of women globally in their reproductive age, is considered to be amongst the most prevalent endrocrinological disorders [1]. It is usually driven by an excessive rise in male androgen hormones in the female body creating a long-term hormonal imbalance that has a detrimental influence on healthy ovarian function and results in

* Corresponding author.
E-mail address: nazrul@cse.mist.ac.bd (M.N. Islam).

https://doi.org/10.1016/j.heliyon.2023.e20524
Received 14 March 2023; Received in revised form 25 September 2023; Accepted 27 September 2023

Available online 5 October 2023

the development of several cysts within the ovary [2]. This condition in women's body shows heterogeneous symptoms that includes anovulation with irregular menstrual cycles, amenorrhea and oligomenorrhea; hyperandrogenism with acne; hirsutism; increasing rate of obesity; and polycystic ovarian structure with numerous cysts inside ovaries [3]. It is a complicated condition that impacts the reproductive, metabolic, and psychological systems in women's body and is strongly linked to several other serious disorders [4]. PCOS is known as the primary cause of anovulatory infertility and is correlated to type 2 diabetes, thyroid problems, increased depression, sexual dissatisfaction, and other conditions which drastically reduces a healthier lifestyle of human being [5]. Moreover, endometrial and ovarian cancer are seem to be more probable to occur in women with PCOS, each of which be catastrophic if not detected timely [6].

However, by using suitable, symptom-focused, long-term, dynamic therapy and medication; the abnormalities in PCOS can be reversed; for which a timely and appropriate diagnosis of PCOS is crucial. A wide range of medical professionals frequently utilize the following three factors popularly known as 'Rotterdam criteria' to determine if a patient has polycystic ovarian syndrome (PCOS): menstrual irregularities, hyperandrogenism, and the presence of multiple cysts on ovarian ultrasound images [7]. But, due to the heterogeneity of the associated symptoms of PCOS, it often becomes ambiguous to the clinicians to detect the condition at an early stage. Furthermore, the most accurate way to diagnose PCOS is to examine the existence of multiple cysts inside ovaries using ultrasonography (USG) imaging technology by expert radiologists [8]. Since this technique also depends on the observer and these images include a lot of noise; clinical assessment, often becomes time-consuming and challenging, with a greater chance for human error. As a consequence, the expensive, device-reliant, and time-consuming procedure of diagnosing PCOS causes many young women who have this critical disorder to go undiagnosed and untreated for extended periods of time; worsening the disease condition.

The automatic identification of PCOS using different computer-assisted approaches has recently gained the attention of numerous researchers as a solution to this issue. Several studies have been conducted to detect PCOS at an early stage in female body employing patients' various clinical, metabolic and diagnostic test records or ovarian ultrasonography (USG) images of patients. However, despite the fact that many research have been published for the automatic detection of PCOS, a tiny proportion of studies have been conducted on thoroughly reviewing these articles. For instance, Deepika et al. [9] conducted a review on the uses of AI for PCOS detection, segmentation, and classification where the author highlighted six studies that used different AI-based methodologies. Thomas et al. [10] reviewed seven studies on the prediction of disease with PCOS morphologies, incorporating hybrid data mining techniques, neural fuzzy logic systems, and the Chi-Square test for feature selection or PCOS classification. Also, Isah et al. [11] conducted a study where they examined nine research articles with computer assisted methods for diagnosing PCOS and detecting follicles in ovarian ultrasound images. However, a scarcity of research has carried out a comprehensive examination of existing literature or a systematic literature review, specifically targeting studies that aim to detect PCOS through various computer-assisted and machine learning based techniques. Moreover, there has been limited exploration by researchers in terms of analyzing the data from various perspectives, such as different themes of data extraction including topological association, data type, scope, research profiles, methodology, etc. Furthermore, few studies have investigated the research gaps present in current studies for predicting PCOS; as well as explored potential avenues for future expansion within this specific field. Yet, rarely any study has specifically focused on analyzing all relevant research pertaining to the diagnosis of PCOS using either symptom-based data or ultrasound (USG) image data. Hence, the novelty of this review study lies in its emphasis on investigating the research gaps identified in previous literature reviews within this domain.

Therefore, in order to facilitate the development of advanced research in this field, this study is focused on conducting a systematic literature review that would analyze the relevant and important articles with all the potential computer-aided solutions for PCOS detection or ovarian cyst identification as well as would investigate the limitations and future research agenda. This study's main goal is to examine the recent research and development perspectives with a particular focus on different computer assisted state-of-art techniques of autonomous PCOS detection through various methodologies; employing either patients' symptoms as clinical feature dataset or ovary ultrasonography (USG) image data. A systematic Literature Review has been conducted to meet the goal of the study for the detail investigation as well as to explore the research gaps and future research agenda that can help significantly to push the knowledge frontier in this domain. The contributions of this review study can be highlighted as:

- Provide a comprehensive view by compiling the state-of-the-art knowledge on digital interventions for predicting PCOS.
- Delve into the topological association, research profiles, data, scopes, methodologies and findings of the existing studies by examining their distinctive insights.
- Provide an overview of the core research findings, current limitations, and future research directions pertaining to the use of computer-aided technologies in autonomously detecting PCOS.

The rest of the article is structured in the following way: Section 2 describes the necessary theoretical background of the study; Section 3 shows the research methodology that has been employed for conducting this SLR; Section 4 elaborately analyzes the selected studies in terms of six data extraction themes and their relevant queries; the summary of the study findings with research gaps and future research recommendations in this domain have been reported in Section 5; and finally the discussion and conclusion of the SLR has been presented in Section 6 with the limitations of the study as well as future plan.

2. Theoretical background

Experts from all over the world are currently placing a strong emphasis on integrating a variety of emerging computerized technologies to diagnose various diseases at an early stage and at a reduced cost which has enhanced significantly the research

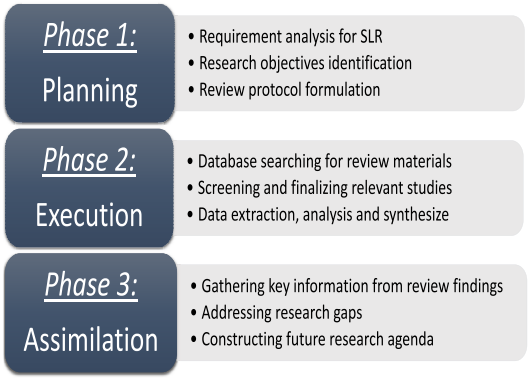

Fig. 1. Outline for systematic literature review.

scopes in healthcare field [12–14]. In recent days an area of artificial intelligence known as machine learning (ML) is getting popular in the field of healthcare analytics that uses a large amount of data for training a machine learning model to find patterns and encourage decision-making [15,16]. ML systems incorporated with big data analytics help to find previously unidentified patterns, stimulating the decision-making procedure where computers are trained to predict or make decision in the same way to humans [17,18]. With the massive expansion of data in healthcare sector, ML is widely employed to analyze electronic health records (EHR) or patient data and create effective clinical decision support systems for different illness diagnosis or forecasting [19,20]. ML techniques are applied to detect various kinds of critical diseases autonomously such as, cardiac anomalies [21], mode of childbirth [22,23], diabetes detection [24,25], Alzheimer's disease diagnosis [26] etc.

Moreover, different types of medical image data are also been used widely to detect the anomalies in patients using various types of methodologies [27]. Researchers worldwide are proposing various predictive image-based techniques using different types of medical images such as Ultrasonography (USG) images, Magnetic resonance imaging (MRI), X-rays etc. Other than machine learning and deep neural network techniques, various types of digital image processing techniques are also widely utilized to extract significant features from medical images and detect anomalies [28,29]. Some areas where medical images are being used for healthcare decision making are: COVID19 detection [30], burn depth severity prediction [31], lung cancer detection [32], brain-tumor detection [33] etc.

Several researchers in recent days have also explored various computer assisted techniques to detect polycystic ovary syndrome or PCOS at an early stage using patient symptom data or medical images. PCOS is a very common disorder among reproductive women which has been proved to be associated with a number of disorders, including cardiovascular disease, type-2 diabetes, hypertension, insulin resistance, dyslipidemia, enhanced Endometrium thickness, ovarian cancer etc. [34]. In practical cases, doctors generally investigate into associated symptoms of PCOS in patients' bodies at an initial stage, such as irregular periods, unwanted facial or body hair, faster hair loss, dark patches on the skin, a high BMI, sudden obesity, and abdominal overweight with a larger hip ratio, etc [35]. Moreover, the PCOS patients suffer from variation in the range of hormonal secretion including the hormones Anti-Müllerian Hormone (AMH), luteinizing hormone (LH), Follicle-stimulating hormone (FSH) and so on; which can be detected using some clinical and pathological tests [36]. Thus, a group of researchers have focused on detecting PCOS automatically based on numbers of patient symptoms which are considered as features, for example, the works conducted by Danaei et al. [37], Nandipati et al. [38] etc. On the other hand, examining ultrasonography (USG) pictures of the ovaries to check for the presence of numerous follicular cysts and then determining whether or not the patient has PCOS is the most common method clinically used by doctors to definitively diagnose PCOS [39]. So, another group of researchers have emphasized on detecting the follicles in ovaries using USG images and thus diagnose PCOS. For example, the studies conducted by Nilofer et al. [40], Dewi et al. [41], Suha et al. [42] etc.

3. Methodology

Literature reviews are generally the initial stage of any research work that provide published knowledge and reliable findings from relevant research on a specific topic [43,44]. Systematic Literature Reviews (SLR) offer a more comprehensive, well-organized summary that follows accepted practices and also outline, assess, and discover the research gaps of the published studies being unbiased in accordance with some study objectives and research questions [45]. From methodological perspective, the SLR in this study has been conducted in three phases following the approaches suggested by Kitchenham et al. [46,47] and Tandon et al. [48]. The outline for the SLR has been shown in Fig. 1.

3.1. Phase 1: planning

The first stage of the study is planning that includes the preparation before executing the review with requirement analysis for conducting the SLR; identifying the research objectives and research questions and then formulating the protocol for conducting SLR.

Table 1
Inclusion and exclusion criteria for SLR.

Inclusion Criteria	Exclusion Criteria
Articles with a specific focus on automated computer assisted techniques for PCOS detection	Articles that mentions PCOS but do not includes computer assisted techniques for PCOS detection
Articles that are available in complete or full-text	Articles published as short surveys, reports, assumptions, notes.
Articles published since 2005 to 2022	Duplicate research articles published in multiple repositories
Research articles published in peer-reviewed journals, conference proceedings or as thesis dissertations	Studies that focus only on clinical diagnosis aspects for PCOS detection
Articles written in English language	Research articles conflicting the objectives of the systematic review

3.1.1. Requirement analysis:

At the requirement analysis phase, the need of performing this SLR has been investigated and therefore necessary background knowledge has been gathered including clinical perspectives of PCOS as well as how computer assisted techniques can be beneficial in PCOS detection. At this stage, it has been observed that numerous research has been carried out that demonstrate how effectively PCOS may be recognized at an early phase through various ML-based techniques using patients' records. However, hardly any study has been carried out where the summary of all these studies has been presented or the findings have been investigated to explore the limitations and future aspects in this research domain. As a result, the requirement analysis step amply demonstrates the need for an SLR in this area.

3.1.2. Research objectives identification:

The core objective of the research is to explore the effectiveness and significance of computer assisted techniques for detecting PCOS; and summarize the findings, limitations and future scopes of the existing studies in a systematic manner. Thus, at this stage of planning the following auxiliary goals have been identified that can help to achieve the primary goal of the study:

- **RO1:** to analyze the purpose, goals, and publication histories of the existing research in this domain;
- **RO2:** to investigate the type of data employed for PCOS identification in relevance to the research's scope;
- **RO3:** to understand different types of computer assisted techniques and evaluation strategies that have been used for autonomous PCOS detection;
- **RO4:** to explore the research findings and challenges to detect PCOS
- **RO5:** lastly, to structure potential future research scopes in autonomous detection of PCOS.

3.1.3. Protocol formulation:

An initial research protocol has been formulated in this stage with inclusion and exclusion criteria for selecting the articles for reviewing. The inclusion and exclusion criteria has been shown in Table 1.

3.2. Phase 2: execution

The execution part of the SLR is where the review has been conducted that involved searching, filtering, and selecting the study materials, after which the required data were extracted, assessed, and synthesized to explore the review outcomes. The activities of execution phase had been performed from February 2022 to July 2022.

3.2.1. Searching, screening and finalizing review materials

For selecting the studies to perform SLR, a systematic guideline known as 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)' [49] has been adapted in this study. The PRISMA flow diagram for selecting the review materials has been shown in Fig. 2. The process consists of four phases: identification, screening, eligibility and included.

At identification phase the articles that have been published in peer-reviewed journals, conference proceedings or as thesis dissertation are searched from different familiar databases like Google scholar, Science Direct, Scopus, IEEE Xplore, ACM Digital Library, Springer Link etc. as well as in some additional sources like website searching, citation tracking etc. using some keywords. The list of keywords that have been used for identification phase is listed in Table 2. Initially total 752 research articles have been found in this way which are then minimized in the screening phase. At screening phase, duplicate records, articles with non English language and the articles that are not available in full-text are eradicated for reviewing that resulted into 151 articles. The next phase is eligibility, where the final article which are eligible for the SLR are selected meticulously considering all the inclusion and exclusion criteria demonstrated in Table 1. Therefore, after conducting all these selecting phases finally 28 articles are included for data synthesis and detail analysis of this study.

3.2.2. Data extraction, analysis and synthesize:

At this phase, data have been extracted, analyzed and synthesized from the selected research articles according to the data extraction theme shown in Fig. 3. Here, six aspects of data extraction themes have been considered to fulfil the research objectives of the SLR. A list of queries is outlined on each theme in order to get the particular type of information needed.

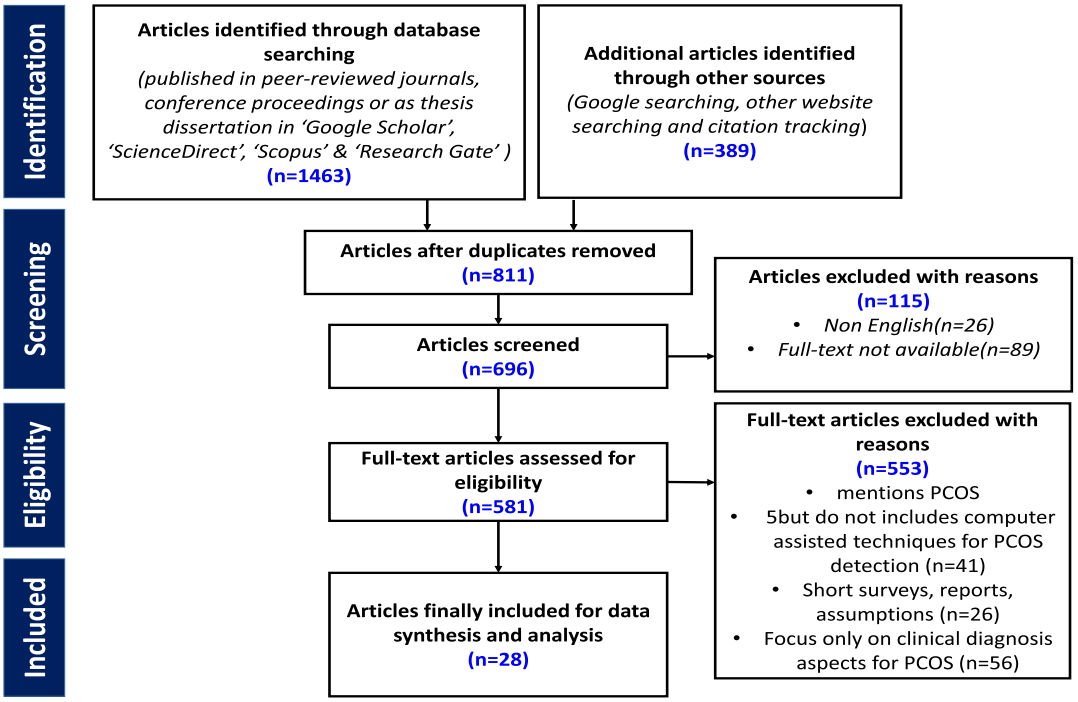

Fig. 2. PRISMA flow diagram for selection of articles.

Table 2
List of keywords utilized for searching in database.

SL No.	Keywords
1	'Polycyctic Ovary Syndrome (PCOS)'
2	'PCOS detection or diagnosis'
3	'PCOS detection with computer assisted' techniques"
4	'Automated PCOS detection techniques'
5	'early screening of PCOS with machine learning techniques'
6	'Feature selection and prediction of PCOS'
7	'Automated Follicle identification in PCOS'

- **Topological Association:** To examine the basic correlation and relevancy between the research literature, this theme investigates two queries:
  - Are the titles of the articles closely linked?
  - Are the keywords in the articles frequent?
- **Research Profiling:** This theme depicts the research publication profiles of the selected articles employing five subqueries:
  - What are the publication types of selected studies?
  - Which year did the studies get published?
  - Which countries are the authors of the articles from?
  - Was funding granted to conduct the research or not?
  - Which publishers published these research articles?
- **Scope of Research:** This theme focuses on the scope and context of the research using the following query:
  - What are the research objectives of the selected studies?
- **Data Profiling:** To investigate the data utilized for the articles according to the research objectives, this theme utilizes three queries:

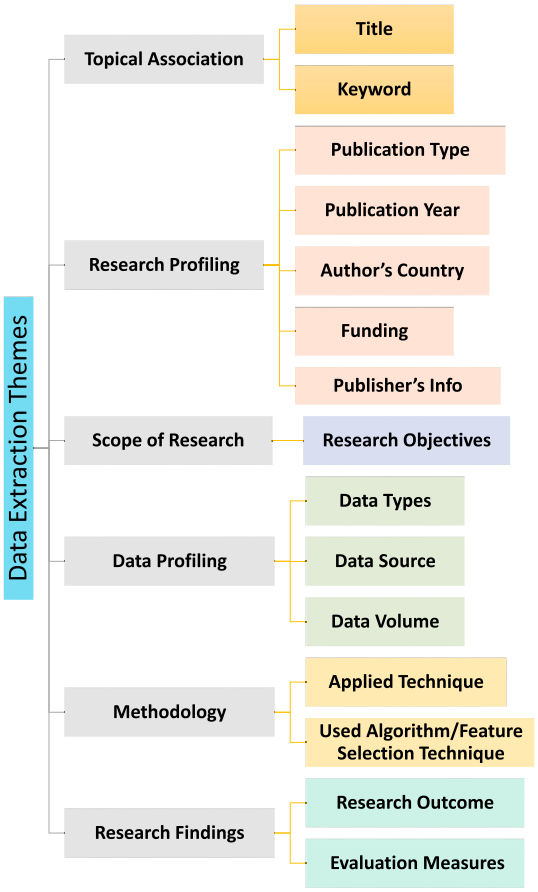

Fig. 3. Data extraction theme for systematic review.

- What type of data has been employed to predict PCOS?
- What are the data sources that have been utilized in the studies?
- What is the volume or size of the dataset that was used?

- **Methodology:** This theme focuses on the research methodology that have been followed to meet the objective of the research employing two questions:

  - What kinds of techniques or technologies have been applied for predicting PCOS in the chosen articles?
  - Which algorithms or feature selection/ extraction techniques have been used to achieve the research objectives?

- **Research Findings:** The key findings from the articles are structured employing the following queries:

  - What evaluation measures have been used to assess the proposed methodology?
  - What are the publications' key research findings?

3.3. Phase 3: assimilation

After conducting the execution step of the SLR, the final stage is assimilation that wraps up the review by reporting the key factors. At this phase, the research results from the selected articles are summarized to present the major findings. Following that, the possible research gaps, limitations and challenges from these investigations are listed. And finally, according to the analysis reports of all the findings the promising future research agenda are explored for autonomous and efficient detection of PCOS employing machine learning and computer assisted advanced methodologies.

4. Data extraction and analysis

Data extraction employing the six themes (see Fig. 3) and their analysis are discussed in this section.

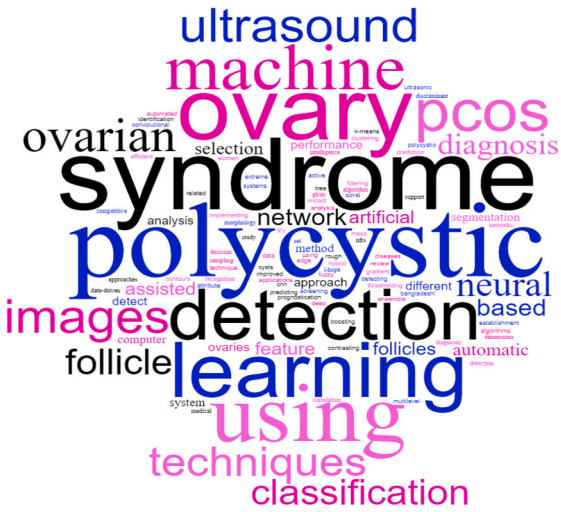

Fig. 4. Wordcloud of the Titles from selected articles.

Fig. 5. Wordcloud of the Keywords from selected articles.

4.1. Topological association

To examine the topological association in terms of the correlation and relevance of the selected articles, 'Wordclouds' have been generated in this study which is considered to be an appealing visualization technique for representing any text analytics quickly [50]. Here, two different word clouds are created, one utilizing all of the article titles (see Fig. 4) and the other using the keywords (see Fig. 5); with an aim to address the two queries under the theme of topological association. The word cloud displays the frequency of the words in proportion to dimensionality that are included in the set of keywords and titles of the articles under evaluation; where the biggest and boldest words tend to be more significant and frequent compared to the smaller, less concentrated words. From the Fig. 4, it is apparent that 'polycystic', 'syndrome', 'ovary', 'learning', 'ultrasound', 'detection', 'machine', 'using' etc. are the most frequently utilized words in the titles of the articles. Also, from the Fig. 5, it is visible that the most frequent words utilized in the keywords of the articles are 'polycystic', 'PCOS', 'ovary', 'syndrome', 'ultrasound', 'follicle', 'feature', 'machine', 'classifier', 'segmentation' etc. Smaller-sized words from the clouds are not frequently utilized in the articles and may have been used as titles or keywords in some individual articles.

4.2. Research profiling

The research profiling of the selected studies has been conducted in five perspectives: publication type, publication year, researcher's country, publisher's name and funding. The publication type includes the information that the selected articles are what type of research articles. From the pie-chart shown in Fig. 6, it is observable that the selected articles are either from any journal

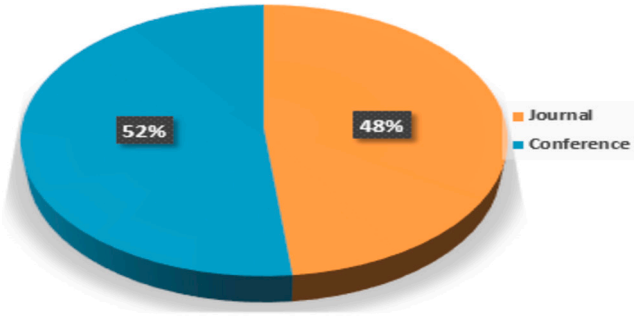

Fig. 6. Publication Type.

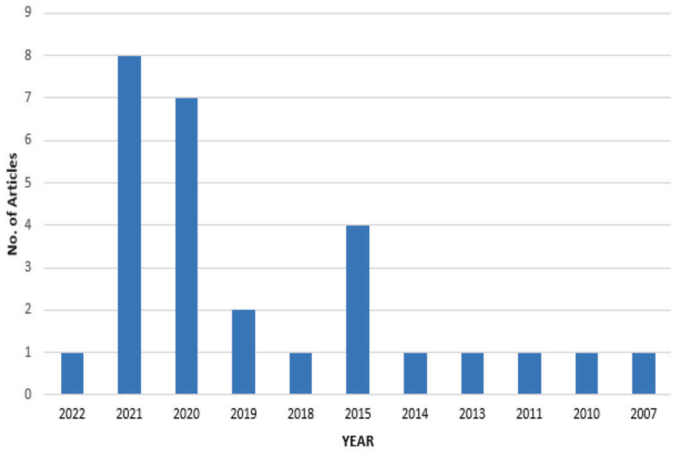

Fig. 7. Publication trend between year 2005 to 2022.

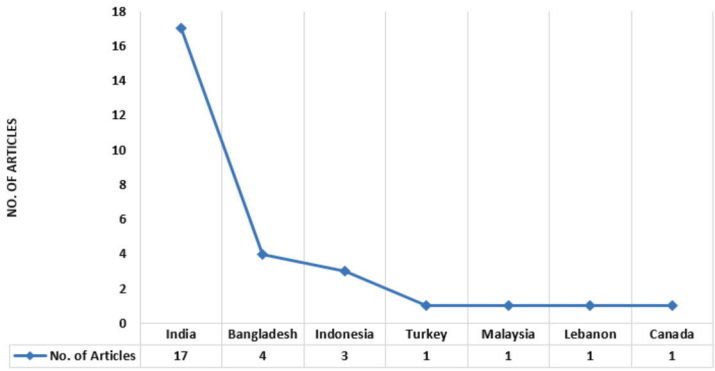

Fig. 8. Number of articles per country where the first author is affiliated.

or from conference proceedings. Here, in 28 selected articles 15 are from journals that covers 52% and 13 are from conference proceedings covering 48% of the publication type.

The chosen research articles of this study have all been published in between the year 2005 to 2022 (up to July) as mentioned in inclusion criteria. Fig. 7 shows the publication trend in between this time frame, which reveals that most of the studies have been conducted in the year 2021. However, from the Fig. 7 it is also observable that, less number of studies had been conducted in this field before 2010 and then the trend of research works is increasing gradually in this domain. For addressing the countries from where the researches in this domain had been conducted, Fig. 8 has been graphically represented. The name of the nation denotes the area of the university that the study's first author is a member of. From the figure it is noticeable that, most of the research works in this domain have been conducted in South-Asian subcontinent where the highest number of studies are from India (17) and then from Bangladesh (4). Other works which have been considered to review in this study are from Indonesia, Turkey, Malaysia, Lebanon and Canada.

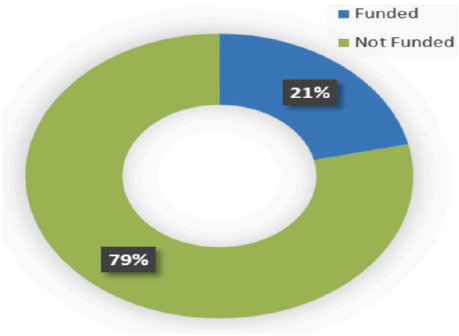

Fig. 9. Funding information.

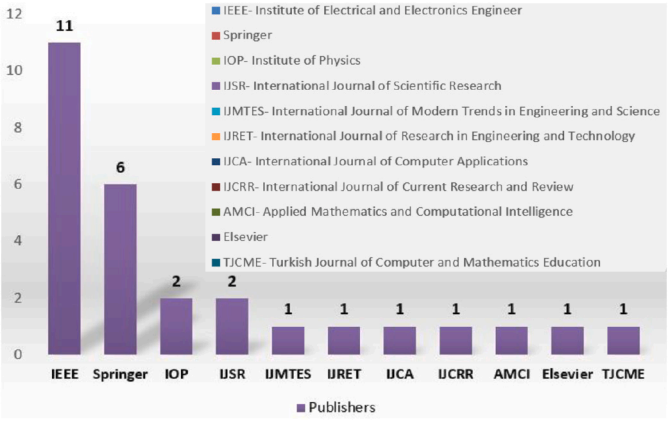

Fig. 10. Number of articles as per Publisher's name.

Table 3
Exploring research scopes of the selected articles.

Category	Research Aim	Reference	Freq
1	To explore the most significant features and then detect PCOS from data-driven patient records	[37,38,51–59]	11 (39%)
2	To detect PCOS from data-driven patient records	[60–62]	3 (11%)
3	To propose an autonomous technique for classifying ovarian ultrasound images in PCOS or non-PCOS classes	[63]	1 (3%)
4	To propose a technique for follicle or cyst segmentation; and then classification into PCOS or non-PCOS classes using ovarian ultrasound images	[40,41,64–69]	8 (29%)
5	To detect the follicle or cyst in ovarian ultrasound images	[70–72]	3 (11%)
6	To classify the type of cyst or follicle using ovarian ultrasound images	[73,74]	2 (7%)

Under the data extraction theme based on research profiling, the following concern is to structure the articles based on its funding options. The donut chart illustrated in Fig. 9 shows that, only 21% (6 out of 28) research works conducted in this domain were funded financially from any organization and most of the studies (79%) were not provided any financial supports to conduct the work. For addressing the final query under this theme, the publishers names are structured as per the number of articles in Fig. 10. From the graphical representation, it is apparent that the highest number of studies that have been conducted to predict PCOS using computer-assisted techniques have been published by IEEE (11 articles), and then by Springer (6 articles). Several other publishers like IOP, Elsevier, IJSR, AMCI etc. have also published number of studies in this research topic.

4.3. Scope of research

With an aim to explore the scopes of the selected research works, Table 3 lists the objectives of the studies as well as their frequencies.

The revised studies were divided into six categories based on their primary study objectives. The first group of studies are those which applied some techniques for feature selection to explore the most significant features from patient records and then using those reduced set of features they applied machine learning classification models to detect PCOS. In this review study, most of the studies (39%) were found with this research objective. For example, Danaei et al. [37] proposed a technique for PCOS diagnosis

Table 4
Data profiling of the selected research articles.

Research Aim Cat	Data Type	Data Source	Data Volume	Ref	Freq
1,2	Patient Records with symptoms & diagnosis results as features	Patient records from hospitals of India (Collected from open source Kaggle data repository [75])	541 Patient records with 44 features	[37,38,51–54, 61,56,62,59]	10
		Online Survey data collected from women of Bangladesh	550 Patient records	[60]	1
		Patients records from Ghosh Dastidar Institute for Fertility Research (GDIFR), Kolkata between March 2010 and April 2011	250 Patient records	[55]	1
		Data from infertility treatment centers at Thrissur, India	540 Patient records, 23 attributes	[58]	1
		Data from GEO dataset-NCBI	303 Patient records, 26 attributes	[57]	1
3,4,5,6	Ovarian Ultrasonography (USG) images from patients	Women's Health Imaging Research Laboratory, (WHIRL) in Saskatoon, Canada	70 images (33 PCOS & 37 normal ovaries)	[65]	1
		Nandhini Sri Diagnostic Center, India.	90 images: normal(30), cystic(25) and PCOS(35)	[66]	1
		Medillab Diagnostics, Gulbarga, India	90 images	[71]	1
		Ovarian mass USG images collected from various hospitals	120 images	[74]	1
		Sardjito Hospital Yogyakarta, Indonesia	100 images	[68]	1
		Various anonymous online open source repositories and websites	Not Specified	[40,41,64,73, 63,67,70,69,72]	9

where the most important 33, 30 and 28 features were selected using several feature selection technique from 42 features and then using those set of features numbers of machine learning classification algorithms were trained and tested to classify among PCOS and non-PCOS patients. Another group of studies (for example the work in [60]) categorized as second categories of research studies as presented in Table 3, did not conduct any feature selection tasks; rather they just applied machine learning classification models employing all the attributes for PCOS detection using data-driven patient records. A limited number of studies have been explored (11%) with such kind of research aim.

Again, the third category of studies in this domain as per the research objective had proposed machine learning based autonomous technique to classify PCOS or non-PCOS patients using ovarian ultrasound images. This category is relatively less explored, as only one study has been found [63]. Following that, the fourth category of research works conducted studies with an aim to initially detect or segment the cysts using various techniques from ovarian ultrasound images; and then employing the extracted features of detected cysts apply machine learning classification methods to classify images into PCOS or non-PCOS category (for example the work conducted by Nilofer et al. [40]). Comparatively a lot of studies (29%) were conducted with this research aim. Moreover, another small group of studies (11%), mentioned as the category five in Table 3, had carried out their research (for example as performed by Raj et al. [70]) with the goal of locating the follicles or cysts as the region of interest in ovarian ultrasound images of PCOS patients. Finally, the studies in the last category of research works (7%) had performed their studies with an aim to classify the type of the detected cysts from the ovarian ultrasound images of PCOS patients as done by Sumathi et al. [73] and Pathak et al. [74]. Thus, from this analysis it is evident that, the researchers have explored the detection of PCOS using computational intelligence in various aspects and aims.

4.4. Data profiling

This data extraction theme focuses on exploring the data profiles of the research article as well as providing a relationship between the research aims and data types of the studies. Table 4 shows the data profiling summary of the selected research articles. The table breaks down the articles into categories based upon several categories of the research purpose while also providing information about the data type, source, volume, and frequency of the articles.

From the table, it is observable that mainly two kinds of data have been utilized to predict PCOS with computer-assisted techniques. The 14 studies with research objective category 1 and 2 (see Table 3) have used patient records as data, where various patient history, symptoms and diagnosis results are represented numerically or categorically as features in the dataset. On the other hand, the other 14 studies with research objective category 3,4,5,6 (see Table 3) have employed ovarian ultrasolonography (USG) images from patients as dataset to predict PCOS. The highest number of studies (10 articles) have employed an open source data from Kaggle repository [75], which consists of the records of 541 patients having 44 attributes collected from various hospitals of India. Other studies have collected data from various hospitals or infertility centers as mentioned in Table 4. However, 9 publications have been discovered which just mentioned anonymous online open source repositories without expressly disclosing their data sources for USG image data.

Table 5
Applied Techniques for the selected studies.

Name of Applied Technology/Technique	Ref	Freq
Machine Learning classification	[37–69], [73], [74]	25
Deep Learning Techniques	[61,40,41,63]	4
Digital Image Processing Techniques	[40–74]	13
Feature Selection Techniques	[37–59]	11

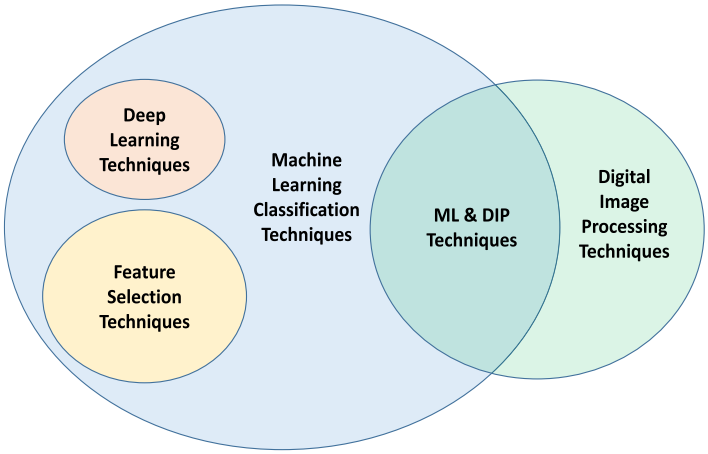

Fig. 11. Ven Diagram of applied research methodologies.

Table 6
Feature selection techniques used in different studies.

Feature selection techniques	Ref
Sequential backward selection (SBS)	[37]
Pearson method	[37]
Random Forest embedded feat. selection	[37]
Recursive Feature Elimination (RFE)	[38,54]
Chi-Square feature selection	[38,56]
Forward Selection & Backward Elimination Technique	[38]
Genetic Algorithm	[51]
Decision Tree	[52]
Filter Method	[53]
Univariate Feature Selection Method	[54]
Statistical Analysis (t-test)	[55]
Analysis of variance (ANOVA) Test	[56,59]
Neural Fuzzy Rough Set (NFRS) & Artificial Neural Network (ANN)	[57]
Correlation based Feature Selection (CFS)	[57]
Principal Component Analysis (PCA)	[57,58]

4.5. Methodology

This data extraction theme focuses on the research methodologies or applied techniques which have been employed in the articles with an aim to autonomously detect PCOS. Table 5 shows the list of major techniques or technologies that have been applied in the articles to achieve the goal. From the table, it is noticeable that among the 28 selected articles, 25 studies have applied machine learning classification techniques. While 3 articles have applied deep learning based methodologies. Again, various kinds of digital Image processing techniques have been applied by 13 articles which utilized USG images for PCOS identification. Moreover, 11 articles have been explored which applied various feature selection techniques to identify the most significant attributes among the attributes in the dataset. The list of several types of feature selection techniques that had been applied in these articles is provided in Table 6. Furthermore, the core applied methodologies are represented through a Ven-diagram in Fig. 11 which depicts that, most of the articles here applied machine learning(ML) techniques and among them four studies applied deep learning techniques and eleven used various feature selection techniques. Moreover, there are some articles which applied digital image processing (DIP) techniques and a few have applied both ML and DIP techniques to achieve the research objectives.

Most of the articles (n = 25) utilized different types of machine learning algorithms for predicting the presence of PCOS employing either the ovarian USG images of patients or their metabolic and clinical attributes. The list of ML algorithms has been provided in Table 7. From the table it is observable that most of the articles utilized Support Vector Machine (n = 12), Random Forest (n = 10)

Table 7
ML algorithms used in different studies.

Algorithms	Ref	Freq
Random Forest Classifier	[37,38,51,52,54,60,56,58,62,66]	10
Multi-Layer Perceptron(MLP)	[37,54,56]	3
K-Nearest Neighbour(KNN)	[38,52–54,60,56,58,59,65,69]	10
Support Vector Machine(SVM)	[38,52,60,56–58,62,59,74,65–67]	12
Gaussian Naive Bayes	[38,52,60,55,57,58,62,66]	8
Decision Tree	[51,54,60,57]	4
Logistic Regression	[53,54,60,55,58,62]	6
Linear discriminant classifier	[65,66]	2
Classification & Regression Trees (CART)	[58,62]	2
Light Gradient Boosting Model(LGBM)	[54]	1
Gradient Boosting(GB)	[54,60]	2
Ensemble Extra Tree	[37,51]	2
Adaptive Boosting(AdaBoost)	[37,38,54,56]	4
Categorical Boosting(CATBoost)	[54]	1
Extreme Gradient Boosting(XGBoost)	[60,56]	2
Voting Hard Ensemble Classifier	[54]	1
Voting soft Ensemble Classifier	[54]	1
Artificial Neural Network(ANN)	[61,57,40]	3
Convolutional Neural Network (CNN)	[41,73,63]	3

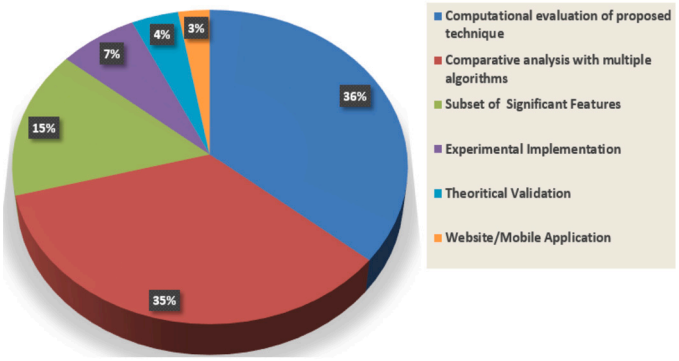

Fig. 12. Varieties of research outcomes.

and K-Nearest Neighbour (n = 10) classification models for training. Different types of ensemble classifiers were also employed by some articles, for example, voting hard and soft ensemble classifier, adaptive boosting, gradient boosting, eXtreme gradient booting, Categorical Boosting classifier etc. Limited number of studies had applied neural network based classification models such as Multi-layer perceptron, Artificial Neural Network, Convolutional Neural Network etc.

Different types of platforms, programming languages and integrated development environment (IDE) were utilized to implement the proposed techniques. For example, Munjal et al. [51] had implemented the feature selection applying genetic algorithm using Waikato Environment for Knowledge Analysis (WEkA) software. Nandipati et al. [38] performed the same methodology in two types of application platforms: Pyhon-Scikit learn and Rapid Miner. However, most of the studies had implemented their proposed techniques with machine learning employing Python programming language based IDE such as, Pycharm, Pydev, Visual Studio, Google Colaboratory etc. Also the articles which have applied digital image processing techniques for follicle segmentation have mostly utilized MATLAB platform for implementation.

4.6. Research findings

The selected studies of this SLR provided several types of research outcome, though all of them had the key goal to detect PCOS through computer-assisted techniques. (See Fig. 12.) The research' findings can be summarized and divided into six groups: subset of significant features for PCOS identification, computational evaluation and validation of the proposed technique, comparative analysis with multiple algorithms or existing methodologies, architectural framework, theoretical validation and website/mobile application. Most of the studies (n = 26) provided a research outcome with computational evaluation and validation of the proposed techniques, where they initially proposed a technique for PCOS detection and then evaluated and validated it through train and test dataset using various performance metrics. For example, Rachana et al. [69] proposed a technique based on image processing and machine learning classification models for PCOS detection; and then conducted a computational evaluation through train and test dataset of ovarian USG images with an aim to explore the efficacy of the proposed methodology. Moreover, a large group of articles (n = 25) provided a comparative analysis between various types of machine learning classifiers with an aim to explore the best performing one

Table 8
Performance metrics used for evaluating the proposed methodologies of the articles.

Metrics	Ref	Freq
Accuracy	[37–41], [73,63,74,65,67,69]	23
Sensitivity (or Recall)	[37–40], [74,65]	16
Specificity	[37,55,58,74,65]	5
Precision	[37–40]	16
F1-Score	[37,51–54,60,56,58,62,40]	10
Confusion Matrix	[51,54,60,56]	4
Area Under the Curve(AUC) Score	[51]	1
Receiver Operating Characteristic(ROC) curve	[62]	1
Positive Predictive Value (PPV)	[55]	1
Loss	[62]	1
Error Rate	[66,71]	2
Comparative evaluation with manual process	[64–66,71,72]	5

or showed a comparison between the proposed computational technique and the manual process of PCOS detection. For example, Gopalakrishnan et al. [66] provided their study findings with a comparative analysis of the manual identification vs. the proposed automated segmentation of follicles for PCOS detection as well as evaluated the error rates.

Moreover, a subset of the most significant features for PCOS detection were suggested as study findings by certain research (n = 11) using the dataset of patients' varied metabolic and clinical symptoms. For example, Danaei et al. [37] applied various types of feature selection techniques such as filter methods, wrapper methods and embedded methods to explore different sets of features selected by different kinds of techniques; and finally suggested a set of features as research outcome which provides best accuracy while applying machine learning models. Another group of studies (n = 5) offered experimental implementation of their proposed methods for PCOS detection or follicle segmentation from ovarian USG images. For example, Lawrence et al. [65] conducted an experimental implementation of their proposed methodology over 70 patients of Women's Health Imaging Research Laboratory (WHIRL) situated in Saskatoon, Canada where they validated the segmentation of the follicles in the ovarian USG images through both expert identification and automated detection process.

A few studies (n = 3) provided a model for web-based or mobile application as research outcome to detect PCOS automatically using machine learning based techniques. For example, Denny et al. [58] suggested an autonomous application model for PCOS diagnosis aid named 'i-Hope' that would be developed based on the attributes selected by Principal Component Analysis (PCA) technique and Random Forest Classification model. Moreover, some studies (n = 2) proposed a technique for PCOS detection and validated it through theoretical explanation. An example of such kind of research outcome can be observed in the work conducted by Hartati et al. [68] where they provided a framework with theoretical and mathematical validation for PCOS identification using ovarian USG images and the researchers hoped to use the suggested method in practice in the future.

However, for evaluating the performances of the proposed techniques, the articles applied different types of performance metrics. The list of the performance parameters with the references and frequencies is presented in Table 8. From the table, it is observable that, most of the articles (n = 25) employed accuracy to measure the efficacy of the proposed technique in result analysis. Precision, recall, F1-scores, specificity, confusion matrix are also utilized by a vast group of articles. Additionally, as PCOS detection is a practical challenge in the field of healthcare analytics, some publications (n = 5) compared the reliability of autonomous detection with the manual identification of PCOS through expert clinicians. Other than these, Area Under the Curve (AUC) score, Receiver Operating Characteristic (ROC) curve, loss calculation, percentage of error rate, positive predictive value (PPV) calculation are also employed to measure the performances of the proposed techniques in few other articles, each with a frequency of 1 or 2.

5. Reporting the review

5.1. Summarizing review findings

This SLR has explored 28 relevant articles which are all been conducted with an aim to detect PCOS in female body autonomously with computer-assisted techniques. For analyzing the articles in a systematic way, the data synthesis was performed using six data exploration themes. The findings from each of the themes can be briefly summarized hereafter:

- In the first theme, the word clouds under topological association reveal that, the titles and keywords of the chosen studies are closely associated with each other with several similar and frequently utilized words where majority of the articles clearly mentions PCOS detection techniques in their titles or keywords.
- The research profiling of the articles clearly indicates that, this area of research has gained almost equivalent publication interest in both journals and conferences from various publishers. The year-wise publication statistics also shows that the interest of the researchers in this domain is increasing day by day. Thus, this analysis indicates that there is a growing emergence with increasing interest by the researchers worldwide in autonomous prediction of PCOS using computer-assisted techniques and machine learning. However, the authors origin statistics indicate that, most of the studies have been conducted in South-Asian subcontinent and majority of them are from India. Also, a few research has acquired funding to conduct such kind of research works.

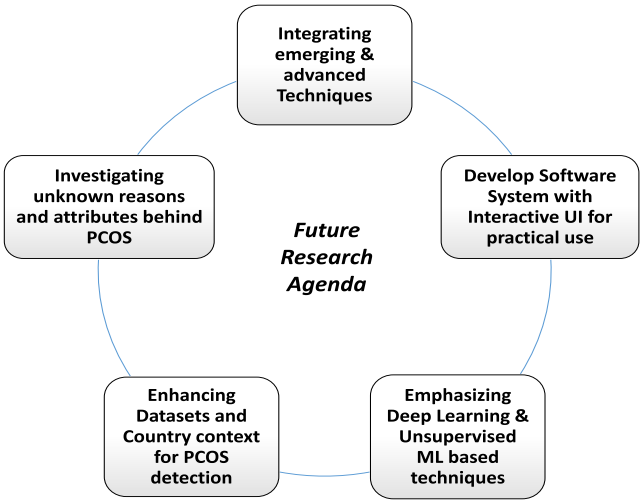

Fig. 13. Outline for future research agenda.

- Investigating the scopes of the selected research articles, it has been explored that, though detecting PCOS using computer-assisted methods is the main goal of all the publications, each one also has certain distinctive research goals. The research objectives of all the articles have been categorized into six groups where 39% studies have the aim to explore the most important features for PCOS detection and then to detect PCOS based on those attributes. And, another large group of studies with 29% articles have the objective to conduct a segmentation of the cysts from ovarian USG images and then classify them into PCOS or non-PCOS criteria.
- The data profiling analysis of the articles indicates that, there are mainly two types of data that have been utilized for PCOS detection: one of which is patients' various symptoms with metabolic and clinical test reports represented as features of the dataset and another type of data is ovarian USG images of patients. However, most of the articles have used dataset which contains the records of Indian women collected from various hospitals or open source repositories.
- The research methodologies of the articles indicate that, there are mainly four types of technologies that have been applied in this domain: machine learning, digital image processing, feature selection techniques and deep learning. Here, most of the studies applied various types of machine learning classification models such as, random forest, KNN, SVM, DT etc. (see Table 7). Also, various types of feature selection techniques have been employed by the articles to explore the most significant features for PCOS detection. A group of articles applied both digital image processing techniques for follicle segmentation and then applied machine learning techniques for PCOS and non-PCOS patient classification. However, only three articles have been found that utilize deep neural networks in this domain, limiting the exploration of deep learning approaches.
- Finally, the findings of selected articles have been grouped into mainly 6 categories where it has been observed that, most of the articles provided computational evaluation and validation as well as comparative analysis of the proposed techniques with other methods as research outcome. Various studies have utilized different types of performance metrics to evaluate the efficacy of their suggested methods. It has also been noticed that, though similar types of classifiers have been employed in multiple articles, each of them gained different performances because of their non-identical methods of implementation and data pre-processing techniques. For example, using the same dataset Danaei et al. [37] acquired 98.89% accuracy whereas Munjal et al. [51] gained 85% accuracy both applying Random Forest classification model for PCOS identification.

5.2. Research gaps and future research recommendations

In the field of PCOS diagnosis utilizing computer assisted techniques, the results of this SLR highlight potential research gaps or constraints that still need to be addressed and also suggests possible future research areas. Following are some suggested future study initiatives that might be taken into account to fill in the research gaps. The possible research agenda are briefly outlined in Fig. 13.

5.2.1. Integrating emerging & advanced techniques

The existing ML-based strategies for PCOS identification can be strengthened by incorporating a variety of emerging and advanced techniques. For example, 'explainable AI or XAI' can be integrated for PCOS detection which contains the potential to explain AI-based decision-making in clear terms to individuals and to a wider range of end users for making it easily interpretable [76]. XAI is gaining much attention in the healthcare analytics as a great tool for clinicians to spread the use of ML-based procedures in practical fields. Moreover, 'Federated Learning or FL' is another cutting-edge method that can be incorporated to detection process of PCOS. FL uses the cloud to prevent security breaches by allowing different data owners to work together to train a model locally employing their own data without sharing data with anyone else [77]. Employing FL techniques in PCOS detection, several hospitals around

the world can participate in developing a robust ML-based model for PCOS prediction without even disclosing their own datasets to others.

5.2.2. Develop software system with interactive UI for practical use

The existing studies in this domain have hardly focused on developing an usable software system with an interactive user interface(UI) [58] so that doctors or patients can utilize it practically for detecting PCOS. Therefore, in addition to performing research on PCOS prediction, ML-based applications can be built in the future and used in real-world scenarios. Additionally, these software systems' usability and effectiveness can be assessed to see how well they enable physicians to make clinical decisions.

5.2.3. Emphasizing deep learning & unsupervised ML based techniques

The articles that have been reviewed in this SLR have all employed supervised machine learning techniques for PCOS detection. However, unsupervised ML techniques like clustering methods can be employed to reveal the patterns of the attributes behind PCOS as well as other useful information. Moreover, a limited numbers of articles have employed deep learning based methodologies which can be explored in a large scale in future. As such, deep learning architecture can also be employed to extract the significant features from the dataset for PCOS detection.

5.2.4. Enhancing datasets and country context for PCOS detection

The SLR findings reveal that, most of the articles have utilized datasets which are from the women of South-Asian subcontinent and specifically from India. But, PCOS is a global problem which is very common in women worldwide. A recent study conducted by Safiri et al. [78] showed that, from 1990 to 2019, experts in 204 nations and territories discovered PCOS to be one of the most prevalent diseases among reproductive women. Therefore, the dataset that are to be employed in the future research in this domain should contain records from various country contexts throughout the world, so that the demographic and geographic impacts of this disease can be well investigated. Additionally, the ovarian USG imaging collection has to be enhanced further with good preservation in a clinical database so that researchers throughout the world may get more use out of it.

5.2.5. Investigating unknown reasons and attributes behind PCOS

All of the research so far in this field has been on the PCOS identification analyzing the patient records. However, the pathogenesis of PCOS is not entirely understood by the clinicians yet, which may be influenced by genetics, lifestyle, and a patient's lack of necessary micro-nutrients [79]. Therefore, it's important to look into the causes of this anomaly as well as any other contributing factors such as environmental impact, socio-economic condition etc. associated with PCOS. Also, the clinicians found PCOS to be linked with various other metabolic and physiological disorders like obesity, insulin resistance (IR), ovarian cancer etc. [80]. The researchers can focus on investigating the association between these anomalies and PCOS using machine learning based techniques.

6. Discussion and conclusion

The existing relevant studies concentrating on the use of various computer-assisted methodologies for PCOS detection have been thoroughly examined in this systematic literature review to explore the state-of-art viewpoints, identify the shortcomings and determine the possible future study areas. The relevant articles have been selected meticulously using PRISMA methodology with distinct inclusion-exclusion criteria. To address the first research objective (R01) the purpose, goals, and publication histories of the existing studies relevant to PCOS detection have been investigated. The data sources, type and attributes utilized in the examined studies with its relevance to research scopes have been investigated to address second research objective (RO2). The SLR also investigated into different computer-assisted methods, algorithms utilized in different studies and evaluation approaches that have been applied to automatically identify PCOS with an aim to fulfil third research objective (RO3). Furthermore, for addressing fourth research objective (RO4) the study has summarized the overall findings from the data synthesis from selected studies as well as the research gap. Finally, some suggestions have been provided concerning possible future study areas in the field of autonomous PCOS detection with enhanced efficiency with a goal to address fifth research objective (RO5).

One drawback of this literature review is that distinct exclusion-inclusion criteria may result in the exclusion of certain relevant and important publications; such as some publications could not be considered because the whole text was not available or they were written in a language other than English. Additionally, some associated articles might not be included because of the fixed set of lists of keywords used to search for articles. The authors of this study intend to undertake a more thorough review in the future, incorporating newer, more pertinent research papers as well as PCOS-related clinical studies.

Therefore, it is obvious that there is a vast array of potential areas for research in this domain, and the findings of this review will undoubtedly aid scholars in understanding the depth and breadth of existing studies, requirements and limitations to conduct more experiments and studies for PCOS detection in future. Additionally, this SLR also will aid the scholars and healthcare professionals to address the usefulness of different computer-assisted techniques, such as machine learning, image processing etc. in real-time decision-making for illness diagnosis at an early stage.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Sayma Alam Suha: Formal analysis, Methodology, Resources, Writing – original draft, Writing – review & editing. **Muhammad Nazrul Islam:** Conceptualization, Formal analysis, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Data used in this research are reported in the article.

References

[1] M. Dapas, A. Dunaif, Deconstructing a syndrome: genomic insights into PCOS causal mechanisms and classification, Endocr. Rev. (2022), https://doi.org/10.1210/endrev/bnac001.
[2] S. Abraham Gnanadass, Y. Divakar Prabhu, A. Valsala Gopalakrishnan, Association of metabolic and inflammatory markers with polycystic ovarian syndrome (pcos): an update, Arch. Gynecol. Obstet. 303 (2021) 631–643.
[3] V. Rodriguez Paris, X.Y.D. Wong, S.M. Solon-Biet, M.C. Edwards, A. Aflatounian, R.B. Gilchrist, S.J. Simpson, D.J. Handelsman, N.O. Kaakoush, K.A. Walters, The interplay between pcos pathology and diet on gut microbiota in a mouse model, Gut Microbes 14 (2022) 2085961.
[4] A.E. Joham, H.J. Teede, Pcos—a metabolic condition with health impacts on women and men, Nat. Rev. Endocrinol. 18 (2022) 197–198.
[5] M. Kałużna, M. Człapka-Matyasik, K. Wachowiak-Ochmańska, J. Moczko, J. Kaczmarek, A. Janicki, K. Piątek, M. Ruchała, K. Ziemnicka, Effect of central obesity and hyperandrogenism on selected inflammatory markers in patients with pcos: a whtr-matched case-control study, J. Clin. Med. 9 (2020) 3024.
[6] B. Meczekalski, G.R. Pérez-Roncero, M.T. López-Baena, P. Chedraui, F.R. Pérez-López, The polycystic ovary syndrome and gynecological cancer risk, Gynecol. Endocrinol. 36 (2020) 289–293.
[7] R. Wang, B.W.J. Mol, The rotterdam criteria for polycystic ovary syndrome: evidence-based criteria?, Hum. Reprod. 32 (2017) 261–264.
[8] S. Bednarska, A. Siejka, The pathogenesis and treatment of polycystic ovary syndrome: what's new?, Adv. Clin. Exp. Med. 26 (2017).
[9] V. Deepika, Applications of artificial intelligence techniques in polycystic ovarian syndrome diagnosis, J. Adv. Res. Technol. Manag. Sci. 1 (2019) 59–63.
[10] N. Thomas, A review on prognosis of pcos using nfrs, hybrid technique and chi-square test, in: Proceedings of National Conference on Advanced Computing and Communication Technology, ISBN 97893-52886-869, 2019.
[11] O.R. Isah, A. Usman, A. Tekanyi, A review on computer assisted follicle detection techniques and polycystic ovarian syndrome (pcos) diagnostic systems, 2015.
[12] T.L. Strome, Healthcare Analytics for Quality and Performance Improvement, John Wiley & Sons, 2013.
[13] N. Imtiaz Khan, T. Mahmud, M. Nazrul Islam, Covid-19 and black fungus: analysis of the public perceptions through machine learning, Eng. Rep. 4 (2022) e12475.
[14] K.T. Shahriar, M.N. Islam, M.M. Anwar, I.H. Sarker, Covid-19 analytics: towards the effect of vaccine brands through analyzing public sentiment of tweets, Inform. Med. Unlocked 31 (2022) 100969.
[15] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, Science 349 (2015) 255–260.
[16] M.N. Islam, K.R. Raiyan, S. Mitra, M.R. Mannan, T. Tasnim, A.O. Putul, A.B. Mandol, Predictis: an iot and machine learning-based system to predict risk level of cardio-vascular diseases, BMC Health Serv. Res. 23 (2023) 171.
[17] M. Ghiasi, Z. Wang, M. Mehrandezh, S. Jalilian, N. Ghadimi, Evolution of smart grids towards the Internet of energy: concept and essential components for deep decarbonisation, IET Smart Grid 6 (2023) 86–102.
[18] M.M. Rahman, N.I. Khan, I.H. Sarker, M. Ahmed, M.N. Islam, Leveraging machine learning to analyze sentiment from Covid-19 tweets: a global perspective, Eng. Rep. 5 (2023) e12572.
[19] R. Bhardwaj, A.R. Nambiar, D. Dutta, A Study of Machine Learning in Healthcare, 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), vol. 2, IEEE, 2017, pp. 236–241.
[20] Z. Hasan, R.R. Khan, W. Rifat, D.S. Dipu, M.N. Islam, I.H. Sarker, Development of a predictive analytic system for chronic kidney disease using ensemble-based machine learning, in: 2021 62nd International Scientific Conference on Information Technology and Management Science of Riga, Technical University (ITMS), IEEE, 2021, pp. 1–6.
[21] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, P. Singh, Prediction of heart disease using a combination of machine learning and deep learning, Comput. Intell. Neurosci. (2021) 2021.
[22] M. Kowsher, N.J. Prottasha, A. Tahabilder, M. Islam, et al., Machine learning based recommendation systems for the mode of childbirth, in: International Conference on Cyber Security and Computer Science, Springer, 2020, pp. 295–306.
[23] M.N. Islam, S.N. Mustafina, T. Mahmud, N.I. Khan, Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda, BMC Pregnancy Childbirth 22 (2022) 1–19.
[24] N.S. Khan, M.H. Muaz, A. Kabir, M.N. Islam, A machine learning-based intelligent system for predicting diabetes, International Journal of Big Data and Analytics in Healthcare 4 (2019) 1–20.
[25] N.S. Khan, M.H. Muaz, A. Kabir, M.N. Islam, Diabetes predicting mhealth application using machine learning, in: 2017 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), IEEE, 2017, pp. 237–240.
[26] M. Tanveer, B. Richhariya, R.U. Khan, A.H. Rashid, P. Khanna, M. Prasad, C. Lin, Machine learning techniques for the diagnosis of Alzheimer's disease: a review, ACM Trans. Multimed. Comput. Commun. Appl. 16 (2020) 1–35.
[27] N. Goel, A. Yadav, B.M. Singh, Medical image processing: a review, in: 2016 Second International Innovative Applications of Computational Intelligence on Power, Energy and Controls with Their Impact on Humanity (CIPECH), 2016, pp. 57–62.
[28] D. Ganguly, S. Chakraborty, M. Balitanas, T.-h. Kim, Medical imaging: a review, in: International Conference on Security-Enriched Urban Computing and Smart Grid, Springer, 2010, pp. 504–516.
[29] K. Suzuki, Overview of deep learning in medical imaging, Radiol. Phys. Technol. 10 (2017) 257–273.
[30] T. Liu, E. Siegel, D. Shen, Deep learning and medical image analysis for Covid-19 diagnosis and prediction, Annu. Rev. Biomed. Eng. 24 (2022).
[31] S.A. Suha, T.F. Sanam, A deep convolutional neural network-based approach for detecting burn severity from skin burn images, Machine Learning with Applications 9 (2022) 100371.
[32] Y. Xie, W.-Y. Meng, R.-Z. Li, Y.-W. Wang, X. Qian, C. Chan, Z.-F. Yu, X.-X. Fan, H.-D. Pan, C. Xie, et al., Early lung cancer diagnostic biomarker discovery by machine learning methods, Transl. Oncol. 14 (2021) 100907.

[33] T. Sadad, A. Rehman, A. Munir, T. Saba, U. Tariq, N. Ayesha, R. Abbasi, Brain tumor detection and multi-classification using advanced deep learning techniques, Microsc. Res. Tech. 84 (2021) 1296–1308.
[34] P. Anagnostis, B.C. Tarlatzis, R.P. Kauffman, Polycystic ovarian syndrome (pcos): long-term metabolic consequences, Metabolism 86 (2018) 33–43.
[35] S. George, A. Alex, Assessment of symptoms and diet intake in young adult with polycystic ovary syndrome (pcos), J. Sci. Res. 65 (2021).
[36] N. Malini, K.R. George, Evaluation of different ranges of lh: Fsh ratios in polycystic ovarian syndrome (pcos)–clinical based case control study, Gen. Comp. Endocrinol. 260 (2018) 51–57.
[37] H. Danaei Mehr, H. Polat, Diagnosis of polycystic ovary syndrome through different machine learning and feature selection techniques, Health Technol. 12 (2022) 137–150.
[38] S.C. Nandipati, C. Ying, K.K. Wah, Polycystic ovarian syndrome (pcos) classification and feature selection by machine learning techniques, Appl. Math. Comput. Intell. 9 (2020) 65–74.
[39] S.J. Pulluparambil, S. Bhat, Medical image processing: detection and prediction of pcos–a systematic literature review, Int. J. Health Sci. Pharm. 5 (2021) 80–98.
[40] N. Nilofer, et al., Follicles classification to detect polycystic ovary syndrome using glcm and novel hybrid machine learning, Turk. J. Comput. Math. Educ. 12 (2021) 1062–1073.
[41] R. Dewi, U. Wisesty, et al., Classification of polycystic ovary based on ultrasound images using competitive neural network, J. Phys. Conf. Ser. 971 (2018) 012005.
[42] S.A. Suha, M.N. Islam, An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image, Sci. Rep. 12 (2022) 1–16.
[43] M. Ghiasi, T. Niknam, Z. Wang, M. Mehrandezh, M. Dehghani, N. Ghadimi, A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: past, present and future, Electr. Power Syst. Res. 215 (2023) 108975.
[44] Y. Xiao, M. Watson, Guidance on conducting a systematic literature review, J. Plan. Educ. Res. 39 (2019) 93–112.
[45] I. Islam, M.N. Islam, Digital intervention to reduce counterfeit and falsified medicines: a systematic review and future research agenda, J. King Saud Univ, Comput. Inf. Sci. (2022).
[46] B. Kitchenham, Procedures for Performing Systematic Reviews, vol. 33, Keele University, Keele, UK, 2004, pp. 1–26.
[47] B. Kitchenham, S. Charters, Guidelines for performing systematic literature reviews in software engineering, 2007.
[48] A. Tandon, A. Dhir, A.N. Islam, M. Mäntymäki, Blockchain in healthcare: a systematic literature review, synthesizing framework and future research agenda, Comput. Ind. 122 (2020) 103290.
[49] A.A. Selçuk, A guide for systematic reviews: prisma, Turk. Arch. Otorhinolaryngol. 57 (2019) 57.
[50] F. Heimerl, S. Lohmann, S. Lange, T. Ertl, Word cloud explorer: text analytics based on word clouds, in: 2014 47th Hawaii International Conference on System Sciences, IEEE, 2014, pp. 1833–1842.
[51] A. Munjal, R. Khandia, B. Gautam, A machine learning approach for selection of polycystic ovarian syndrome (pcos) attributes and comparing different classifier performance with the help of weka and pycaret, Int. J. Sci. Res. (2020) 59–63.
[52] A.S. Prapty, T.T. Shitu, An efficient decision tree establishment and performance analysis with different machine learning approaches on polycystic ovary syndrome, in: 2020 23rd International Conference on Computer and Information Technology (ICCIT), IEEE, 2020, pp. 1–5.
[53] N. Tanwani, Detecting pcos using machine learning, Int. J. Mod. Trends Eng. Sci. 7 (2020) 1–20.
[54] S. Bharati, P. Podder, M. Mondal, V. Surya Prasath, N. Gandhi, Ensemble learning for data-driven diagnosis of polycystic ovary syndrome, in: International Conference on Intelligent Systems Design and Applications, Springer, 2022, pp. 1250–1259.
[55] P. Mehrotra, J. Chatterjee, C. Chakraborty, B. Ghoshdastidar, S. Ghoshdastidar, Automated screening of polycystic ovary syndrome using machine learning techniques, in: 2011 Annual IEEE India Conference, IEEE, 2011, pp. 1–5.
[56] M.S.K. Inan, R.E. Ulfath, F.I. Alam, F.K. Bappee, R. Hasan, Improved sampling and feature selection to support extreme gradient boosting for pcos diagnosis, in: 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), IEEE, 2021, pp. 1046–1050.
[57] K. Meena, M. Manimekalai, S. Rethinavalli, Correlation of artificial neural network classification and nfrs attribute filtering algorithm for pcos data, Int. J. Res. Eng. Technol. 4 (2015) 519–524.
[58] A. Denny, A. Raj, A. Ashok, C.M. Ram, R. George, I-hope: detection and prediction system for polycystic ovary syndrome (pcos) using machine learning techniques, in: TENCON 2019-2019 IEEE Region 10 Conference (TENCON), IEEE, 2019, pp. 673–678.
[59] Y.A.A. Adla, D.G. Raydan, M.-Z.J. Charaf, R.A. Saad, J. Nasreddine, M.O. Diab, Automated detection of polycystic ovary syndrome using machine learning techniques, in: 2021 Sixth International Conference on Advances in Biomedical Engineering (ICABME), IEEE, 2021, pp. 208–212.
[60] N. Nabi, S. Islam, S.A. Khushbu, A.K.M. Masum, Machine learning approach: detecting polycystic ovary syndrome & it's impact on Bangladeshi women, in: 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE, 2021, pp. 1–7.
[61] R. Boomidevi, S. Usha, Performance analysıs of polycystıc ovary syndrome (pcos) detectıon system usıng neural network approach, in: Data Engineering and Communication Technology, Springer, 2021, pp. 449–459.
[62] M.M. Hassan, T. Mirza, Comparative analysis of machine learning algorithms in diagnosis of polycystic ovarian syndrome, Int. J. Comput. Appl. 975 (2020) 8887.
[63] V. B, R. Yalavarthi, V. K, Detection of polycystic ovarian syndrome using convolutional neural networks, Int. J. Curr. Res. Rev. 13 (2021) 155–159, https://doi.org/10.31782/IJCRR.2021.13630.
[64] B. Padmapriya, T. Kesavamurthy, Diagnostic tool for pcos classification, in: 7th WACBE World Congress on Bioengineering 2015, Springer, 2015, pp. 182–185.
[65] M.J. Lawrence, M.G. Eramian, R.A. Pierson, E. Neufeld, Computer assisted detection of polycystic ovary morphology in ultrasound images, in: Fourth Canadian Conference on Computer and Robot Vision (CRV'07), IEEE, 2007, pp. 105–112.
[66] C. Gopalakrishnan, M. Iyapparaja, Multilevel thresholding based follicle detection and classification of polycystic ovary syndrome from the ultrasound images using machine learning, Int. J. Syst. Assur. Eng. Manag. (2021) 1–8.
[67] B. Purnama, U.N. Wisesti, F. Nhita, A. Gayatri, T. Mutiah, et al., A classification of polycystic ovary syndrome based on follicle detection of ultrasound images, in: 2015 3rd International Conference on Information and Communication Technology (ICoICT), IEEE, 2015, pp. 396–401.
[68] S. Hartati, A. Musdholifah, et al., Machine learning assisted medical diagnosis for segmentation of follicle in ovary ultrasound, in: International Conference on Soft Computing in Data Science, Springer, 2019, pp. 71–80.
[69] B. Rachana, T. Priyanka, K. Sahana, T. Supritha, B. Parameshachari, R. Sunitha, Detection of polycystic ovarian syndrome using follicle recognition technique, Global Trans. Proc. 2 (2021) 304–308.
[70] A. Raj, Detection of cysts in ultrasonic images of ovary, Int. J. Sci. Res. 2 (2013) 185–189.
[71] P. Hiremath, J.R. Tegnoor, Follicle detection in ultrasound images of ovaries using active contours method, in: 2010 International Conference on Signal and Image Processing, IEEE, 2010, pp. 286–291.
[72] V. Kiruthika, M. Ramya, Automatic segmentation of ovarian follicle using k-means clustering, in: 2014 Fifth International Conference on Signal and Image Processing, IEEE, 2014, pp. 137–141.
[73] M. Sumathi, P. Chitra, R.S. Prabha, K. Srilatha, Study and detection of pcos related diseases using cnn, IOP Conf. Ser., Mater. Sci. Eng. 1070 (2021) 012062.
[74] H. Pathak, V. Kulkarni, Identification of ovarian mass through ultrasound images using machine learning techniques, in: 2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), IEEE, 2015, pp. 137–140.
[75] S. Vedpathak, Pcos dataset, https://www.kaggle.com/shreyasvedpathak/pcos-dataset, 2020.

[76] Y. Zhang, Y. Weng, J. Lund, Applications of explainable artificial intelligence in diagnosis and surgery, Diagnostics 12 (2022) 237.
[77] M.M. Rahman, D. Kundu, S.A. Suha, U.R. Siddiqi, S.K. Dey, Hospital patients' length of stay prediction: a federated learning approach, J. King Saud Univ, Comput. Inf. Sci. (2022), https://doi.org/10.1016/j.jksuci.2022.07.006, https://www.sciencedirect.com/science/article/pii/S1319157822002336.
[78] S. Safiri, M. Noori, S.A. Nejadghaderi, N. Karamzad, K. Carson-Chahhoud, M.J. Sullman, G.S. Collins, A.-A. Kolahi, J. Avery, Prevalence, incidence and years lived with disability due to polycystic ovary syndrome in 204 countries and territories, 1990–2019, Hum. Reprod. (2022).
[79] G. Tefagh, M. Payab, M. Qorbani, F. Sharifi, Y. Sharifi, M.S. Ebrahimnegad Shirvani, F. Pourghazi, R. Atlasi, Z. Shadman, N. Rezaei, et al., Effect of vitamin e supplementation on cardiometabolic risk factors, inflammatory and oxidative markers and hormonal functions in pcos (polycystic ovary syndrome): a systematic review and meta-analysis, Sci. Rep. 12 (2022) 1–16.
[80] R.A. Condorelli, A.E. Calogero, M. Di Mauro, S. La Vignera, Pcos and diabetes mellitus: from insulin resistance to altered beta pancreatic function, a link in evolution, 2017.