Blood Reviews 61 (2023) 101102

Contents lists available at ScienceDirect

Blood Reviews

journal homepage: www.elsevier.com/locate/issn/0268960X

Review

Artificial intelligence in sickle disease

Ahmed Adel Elsabagh [a,*], Mohamed Elhadary [a], Basel Elsayed [a], Amgad Mohamed Elshoeibi [a], Khaled Ferih [a], Rasha Kaddoura [b], Salam Alkindi [c], Awni Alshurafa [d], Mona Alrasheed [e], Abdullah Alzayed [f], Abdulrahman Al-Abdulmalek [g], Jaffer Abduljabber Altooq [h], Mohamed Yassin [i,*]

[a] College of Medicine, QU Health, Qatar University, Doha, Qatar
[b] Pharmacy Department, Heart Hospital, Hamad Medical Corporation (HMC), Doha, Qatar
[c] Professor of Hematology, Sultan Qaboos University, Oman
[d] Department of Hematology, Hamad Medical Corporation (HMC), Doha, Qatar
[e] Hematology Unit, Department of Medicine, Aladnan Hospital, Ministry of Health, Kuwait
[f] Qatif Central Hospital, Saudi Arabia
[g] Internal Medicine Department, Montreal General Hospital, Montreal, Canada
[h] Internal Medicine Department, Bahrain Sulaymaniyah Hospital, Bahrain
[i] Hematology Section, Medical Oncology, National Center for Cancer Care and Research (NCCCR), Hamad Medical Corporation (HMC), Doha, Qatar

ARTICLE INFO

Keywords:
Artificial intelligence
Sickle cell
Machine learning
Convolutional neural networks
Hemoglobinopathies

ABSTRACT

Artificial intelligence (AI) is rapidly becoming an established arm in medical sciences and clinical practice in numerous medical fields. Its implications have been rising and are being widely used in research, diagnostics, and treatment options for many pathologies, including sickle cell disease (SCD). AI has started new ways to improve risk stratification and diagnosing SCD complications early, allowing rapid intervention and reallocation of resources to high-risk patients. We reviewed the literature for established and new AI applications that may enhance management of SCD through advancements in diagnosing SCD and its complications, risk stratification, and the effect of AI in establishing an individualized approach in managing SCD patients in the future. **Aim**: to review the benefits and drawbacks of resources utilizing AI in clinical practice for improving the management for SCD cases.

1. Introduction

Artificial intelligence (AI) or machine intelligence as a term was coined in 1956 [1]. As computing power continues to grow, AI is exerting an immense effect on today's society, particularly in medical practice [2]. AI in simple language means transmitting human cognitive ability to machines [2]. As a programmed machine, AI can learn and distinguish patterns and correlations between inputs and outputs, then use this knowledge for decision-making on the new input data [1]. AI utilizes computational networks, i.e., neural networks, that emulate a biological nervous system. Based on functionality and capability, AI can be grouped into types 1 and 2. Type 1 is based on capabilities. This describes the AI method according to its ability to perform complex tasks, learn from previous input, and take appropriate decisions. Type 2 is focused on functionality. This assesses the ability to exhibit human features like responding to human emotions, holding conversations, and being self-aware. An example of that is robots that display human-like features and functions [3].

AI can be actualized through machine learning and deep learning. Machine learning is a subfield of AI and deep learning is a specific subset of machine learning with a focus on deep neural networks [1]. Briefly, machine learning intends to develop algorithms that process the input data to detect patterns, through statistical analysis, and make inferences. Subsequently, the data sets are utilized to train the machine. Deep learning evolves from machine learning algorithms through the utilization of hierarchical levels of neural networks, that obtain features from raw input [1,3]. Many reviews are now emerging with interest in Artificial Intelligence applications in certain hematological diseases [4,5]. The current clinical use of AI spans numerous specialties such as cardiology, orthopedics, rheumatology, oncology, and hematology [2].

* Corresponding authors.
E-mail addresses: ae1802661@qu.edu.qa (A.A. Elsabagh), ae1802661@qu.edu.qa (M. Yassin).

https://doi.org/10.1016/j.blre.2023.101102

Available online 8 June 2023

Sickle cell disease (SCD) is the most common inherited hematologic disease [2,6,7]. According to the Centers for Disease Control and Prevention (CDC), SCD is prevalent in sub-Saharan Africa, Spanish-speaking regions in the Western Hemisphere, Saudi Arabia, India, and Mediterranean countries [8]. In SCD, sickle hemoglobin (HbS) replaces the normal healthy healthy hemoglobin (HbA). HbS has a life span of ten to twenty days compared to the life span of about 120 days in HbA [9]. Hemoglobin SS (HbSS) is the most common genotype in SCD, which indicates the presence of two copies of the HbS gene. There are several other heterozygous genotypes such as sickle cell plus thalassemia (HbS + thal) in which an individual carries the HbS gene in addition to one or more thalassemia genes. Hemoglobin sickle cell disease (HbSC) genotype may also occur as a patient inherits a copy of the HbS gene and a copy of HbC gene. [10,11]. The red blood cells have a sickle shape because of the process of hemoglobin polymerization of the deoxygenated molecule with HbS. The cell morphology is used to classify patients' disease states [9]. The manifestations of SCD include vaso-occlusive crises, sepsis, ischemic strokes, acute chest syndromes, and aplastic crises. Painful vaso-occlusive crises are widely prevalent and are in fact the most common cause of hospitalizations in SCD patients. In vaso-occlusive episodes, sickle red blood cells stick to white blood cells that adhere to vessel endothelium, leading to microvascular occlusion, hence resulting in tissue ischemia and pain. This may lead to complications like arthritis, retinopathy, renal failure, and strokes according to the site of occlusion. Aplastic crises usually occur due to an infectious cause that leads to destruction of red blood cell precursors and a significant reduction in red blood cell formation [12–15]. The core management of patients with SCD reduces the risk of complications with the use of hydroxyurea. Although red blood cell transfusion reduces the burden of sickled cells, it is associated with iron overload [16]. Although the genotype of SCD is widely known, disease manifestations and the phenotypical pictures vary widely among patients. It is caused by a point mutation of a single base pair in the beta-globin gene, resulting in the change of amino acid valine to glutamic acid when producing the beta-globin chain [8]. SCD is also the world's most common monogenic disease. The objective of this review is to discuss the utilization of AI for the diagnosis of the disease and its complications, treatments, and risk stratification for SCD patients. It is important to note the increase in machine-readable data that has been published recently. The Sickle Cell Ontology (SCDO), a comprehensive domain that describes SCD concepts and terminology, is now represented in machine-readable formats. This means that such data may be used to improve the quality of results yielded in diagnostics, research, prediction of outcomes, and optimization of SCD treatment [17].

2. Materials and methods

2.1. Literature search strategy

An electronic literature search was performed on 1 November 2022 using PubMed to identify primary literature. No language or date or language restrictions were applied to the search strategy. The search strategy was also transferred to other databases using Polyglot translator, them being Cochrane Library, Web of Science, Embase, and Scopus. The following terms were used and combined using Boolean operators 'AND/OR': 'artificial intelligence', 'machine learning', 'deep learning', 'convolutional neural network', 'sickle cell', 'diagnostics', 'pathomics', 'radiomics', and 'radiogenomics'. The references of the identified studies, review articles, systematic reviews, and meta-analyses were manually screened to identify additional studies. The identified studies were grouped according to their objectives such as diagnosis of the disease, its complications, and determining the risk of patient clusters. The role of personalized medicine is also discussed, and papers were then critically appraised. Removal of duplicates was then done. All titles and abstracts were independently screened by the authors.

2.2. Inclusion criteria

Research articles were included if they met the following criteria, 1) The authors used a technique or method that relies on the usage of Artificial Intelligence to function, 2) The research showed conclusions regarding the reliability or accuracy of using such method, 3) Reports results useful to assess in the diagnostics or treatment of sickle cell diseases.

2.3. Role of AI in diagnostics

The challenge faced in the detection of SCD is that red blood cells have heterogenous shapes and may overlap in images. The traditional method used involves using microscopy to identify and count cells and outliers. There are several additional limitations, too. A recent publication by Douglass et al. investigated the use of a lensless encoding biosensor that could differentiate between sickle and healthy red blood cells. The samples collected and used were wet mounts of whole blood samples from human donors. The single random phase encoding (SRPE) system was able to detect the diseased samples. This system is a low-cost, lenseless approach that allowed the authors to identify cells. To elaborate, the SRPE system exposes the sample to a laser light source. Subsequently, the laser is tuned by intracellular structures. The light then travels to a diffuser that obtains details about spatial frequency. The intensity of the pattern is then recorded by a CMOS sensor. This offers a low-cost method for less developed healthcare systems to identify and diagnose the disease in a faster way. Another beneficial aspect was the ability of the developed system to detect differences and detect diseases other than the sickle red blood cell. However, a drawback was that using some crops, or images, was limited by the sensor dimensions. In addition, more data should be captured and used in order to further increase the accuracy of the classifications made. This reduced the ability to use some crops which were one-dimensional. The system that was developed had an accuracy of 88.7%. This is done by using a convolutional neural network, which is a network architecture for deep learning to recognize patterns, objects, or scenes. This was utilized in a SRPE system [18].

An interesting tool was also developed to be used on smartphones which helps aid in the diagnosis of SCD. That is a 3D-printed attachment that uses an RBC sample suspended in a paramagnetic medium and sodium metabisulfite, then differentiates sickle and healthy red blood cells using its magnets. This is used with a cell phone application that helps interpret those results. The steps are taken to get the results usually take under ten minutes to perform. Nonetheless, its limitation is that it could only be used to detect homozygous HbSS genotypes, which makes it unattractive in clinical practice. However, it could start newer waves of innovation to create a smartphone attachment to further differentiate different sickle cell disease subtypes. The developed methods so far still need specific human expertise, which is both costly and labor-intensive. The classifications are also still limited to the accuracy of the classifiers and their performance to make a diagnosis. Hence, more accurate classifications are also required which would include all patterns and types of RBCs in SCD patients [19]. On the other hand, the benefit gained from using deep convolutional neural networks (dCNNs) is the ability to test over a thousand RBCs in a few seconds. This has been outlined in a paper by Xu et al., in which they suggested the idea of using a convolutional neural network as it showed great performance even in deoxygenated datasets. In this method, the authors first extracted blood samples and performed RBC extraction. They then separated touching RBCs and applied RBC patch-size normalization methods to unify the RBC patches into a certain size. They then used their deep CNN system to extract information about the RBC sample. The main classification was based on the shape or type RBCs, in which CNN classified them into discocytes, echinocytes, elongated, granular, oval, reticulocyte, sickle, or stomatocytes. This may also be improved as the system is better trained. In addition, this will enable the creation of multiple

classifications in patients with heterogenous blood cells. They suggest that this method may also be used later to classify the severity of SCD using their framework. Moreover, new clustering techniques are also needed to cluster the cells automatically [20].

González-Hidalgo et al. demonstrated a method to analyze erythrocyte shape and quantify the number of healthy and abnormal cells using an algorithm that has the ability to accurately and efficiently find overlapping clusters and minimize the noise in images. They have done that by using ellipse adjustments to the input as well as designing a system that only detects remarkable points. A positive remark was that they verified this system by testing its efficacy with real images, artificial images, and synthetic images. Their results showed an efficiency of 98% or greater in all those image types, with real images showing the best results. In addition, the methods used were all automated. However, diagnosis still may not be entirely dependent on this method and should be confirmed via additional analyses. In order to validate their results, they held five experiments using real, synthetic, and artificial images. When comparing their method to others, they had the highest efficacy as they minimized mistakes in object detection with an efficacy of 100%. They also had impressive efficiency in detecting all different shapes including elongated and circular objects. When comparing their method to another study that used the ellipse fitting technique and concavity detection, they had a significantly higher efficacy, and no cells were detected as sickle or normal wrongly. Hence, this method would be perfect to count the number of normal and deformed cells [21]. Active Appearance Models (AAM) were another method for segmentation to separate erythrocytes from the background in a precise manner. This set of models can help distinguish cells from the background noise. AAM may also describe the shape and texture of cells accurately by applying a shape variation and texture variation model. This ultimately leads to better segmentation of RBCs to give rise to more accurate analyses to be performed. Another benefit is that since AAM uses closed form of boundary, performing boundary tracking to count cells would be unnecessary [22]. Another recent paper by Darrin et al. uses a combination of convolutional and recurrent neural network machine learning algorithms to analyze certain red blood cell videos in SCD patients. Using this information, they deduced that this technique could be used to assess the red blood cell dynamics and differentiate between normal and sickled cells with high accuracy. They trained their network by providing a large number of videos from a predefined dataset to allow the network to identify features to differentiate between the different red blood cell types. They suggest that the study may be used to monitor the disease and development of new treatments by identification of certain therapeutic targets and monitoring the effects of treatment on cell morphology. These results prove the effectiveness of AI methods in determining biological dynamics and the uses that could be gained from these variables [23].

Other clustering methods were used like k-means and DBSCAN, which are methods that aid in automatic cellular classification. DBSCAN is one of the most widely used and cited clustering algorithms in scientific literature. Simply, it works by grouping points that are closer and packed together in a group and identifies outliers that are in low-density areas and are farther than any other points. On the other hand, k-means is a centroid-based clustering algorithm as clusters are assigned according to their proximity to a certain mean (the centroid). It then identifies the number of clusters (K) found in the dataset classification [24]. Classification techniques also included some that used deep convolutional neural networks with Hep-2 cells [25]. Nonetheless, all the aforementioned techniques share some disadvantages including the need for human expertise and could be time consuming [20]. Alagu S et al. used images from a public database called erythrocyte IDB to extract their features and use that information to build a system that is accurate in detecting SCD and differentiating them from healthy cells. This might also help in the detection of sickle cell SCD-like disorders. After they optimized InceptionV3, an already established system, they had results that showed a sensitivity of 98% and specificity of 99%. Inception V3 is an image-recognition model that helps in the recognition of different components in an image. The authors showed improvement of the results when introducing Multi-Objective Binary Grey Wolf Optimization (MO-BGWO), which is a feature selection tool, in addition to KNN and SVM, which are machine learning algorithms used for classification of images. This might help in early detection of sickle cell disease, which would prevent further complications and may improve disease course [26].

2.4. Role of AI in diagnosing complications

A common complication of SCD is proliferative sick cell retinopathy (PSR). This is a condition in which retinal microvasculature is occluded causing ischemia, which in turn leads to proliferation and ultimately leads to retinal detachment [27,28]. Cai S et al. reviewed the benefits of using AI to interpret retinal imaging in a faster and more easily accessible manner when compared to human expertise. The study's limitations were due to its being a single-institution design using a limited number of images. The population studied, however, was reflective of the population with a higher risk to develop SCD, as they included mostly patients of Black or African descent [27]. The study showed a 97.4% sensitivity and 97% specificity for AI interpreting retinal photographs compared to retinal specialists. AI was used in another study to create a retinal vessel mask on fluorescein angiography (FA) images, which are usually used for staging SCD. This helps compute qualities including the length and area of the vessel in addition to the ischemic index. However, there remain several challenges that hinder the possibility to implement those techniques in clinical practice. Generalizability is a challenge as these systems and CNNs developed must be tested on the target population. This is due to the discrepancy in the images acquired in different health systems around the world. Hence, more open access datasets are required and could be very helpful. There is also a medicolegal aspect that must be considered when relying on these diagnoses [29]. Many studies have also shown that in the future, using several types of imaging to be interpreted by AI and deep learning could further enhance current performance [30]. As a next step, both staging and interpretation of more data may help personalize and determine each individual's risks [31,32].

Machine learning methods may also be able to predict organ dysfunction occurring in patients with SCD. A multilayer perceptron model was used in a study done by Mohammed A et al. In this study, they tried to test the accuracy of predictions made by machine learning regarding organ dysfunction in ICU admitted SCD patients. It showed accurate predictions for organ failure with 96% sensitivity and 98% specificity [33], which demonstrates more promising results by machine learning techniques. Their model was also able to predict organ dysfunction up to 6 h before onset. Another study by Hankins JS et al. showed that significantly enhanced clinical outcomes would be present if the applications were implemented on the levels of both the providers and the patients. For instance, enhancing both hydroxyurea prescribing behaviors and compliance by patients simultaneously would further improve outcomes [34].They used physiological data including blood pressure and respiratory and heart rates. This could be beneficial since early detection of the occurrence of organ failure may lead to earlier treatment and improved outcomes [35]. This can be enhanced further by creating models that may predict the organ system that has the highest risk of failing in patients with SCD.

AI has also been used as systems in wearable devices to detect and diagnose painful crises as early as possible. However, a common limitation in these types of studies is the great difficulty in accurate identification and diagnosis of vaso-occlusive crises. Since SCD vaso-occlusive crises are related to peripheral vasoconstriction due to triggers like obstructive sleep apnea, a machine learning algorithm described in a paper by Ji Y et al. was developed to use non-invasive measurements to predict the occurrence vaso-occlusive crises. These measurements include finger photoplethysmogram (PPG) and heart rate. The aim of the

study was to explore the possibility of having a wearable device that detects susceptibility of each SCD patient to experiencing more painful vaso-occlusive crises. To elaborate, PPG would detect the severity of vasoconstriction during sleep, which is associated with higher frequency of vaso-occlusive episodes. They used this association to build a tool that uses machine learning to predict the frequency of the episodes based on PPG findings. This might help clinicians plan long-term therapy for SCD patients and anticipate painful episodes using non-invasive mesures [36]. Other studies have also shown that current wearable technology has the ability to report objective data that can be used to predict a patient's pain score. In one study, researchers used the physiologic data recorded on these devices and used machine learning to score the pain and to create a prediction model for each patient. The study showed that the wearable device that was used recorded physiologic markers during the occlusive episode that could be matched with pain scores and used in a machine learning model that may predict patient pain scores in future patients. Due to the opioid crisis we are currently facing, this objective measure of pain could be of great use to evaluate the necessity of adding opioid doses in ptients experiencing painful crises [37,38]. Other technological apparatuses are being utilized to spread helpful information to patients and recommend healthy lifestyles and health advice catered to SCD patients as a way to promote the prevention of further complications [39]. The benefits of digital health technologies are still not fully utilized. It has been found to be of great importance to include patients' opinions and suggestions when trying to create technological devices and applications that may compute to the patient's day-to-day needs and provide suitable individualized advice to patients based on their information [40]. Some mobile health applications are attempting to improve outcomes by another method which is improving hydroxyurea prescribing behavior in physicians as well as compliance with the medication in patients [41,42]. In many cases, there are also low usage rates of these applications due to insufficient infrastructure, technology gaps, inability to use the features continuously and persistently, and poor interface. With the clear evidence that such digital health interventions are associated with improved patient outcomes, an effort must be made to overcome those obstacles and increase access to these applications.

A study by Alambo A et al. tried using machine learning to classify SCD related pain by creating two classifications: pain relevance and pain change. They applied different methods of machine learning to patient notes in order to classify the pain described in these notes in terms of relevance and change. The pain relevance classifier categorized notes into either pain relevant or irrelevant clinical notes. The pain change classifier then categorized the pain relevant notes into one of the following categories: 1) pain increase, 2) pain uncertain, 3) pain unchanged, and 4) pain increased. The rationale behind the study is to test those machine deep learning techniques and their abilities to predict pain in the future using only clinical notes. This is an attempt to later develop systems that may predict the progression and patterns of pain in patients with SCD [43]. Another study by Yang F et al. also showed a promising ability to predict pain scores in an intra-individual and inter-individual level analysis. Intra-individual analysis is used when patients have sufficient recorded data that enables them to have an individualized model. The inter-individual model, however, is a general model that was created using data from all other patients, so it could be applied to any new patients. The intra-individual level showed a higher level of accuracy, but both were close. The objective is to create an objective way of scoring and predicting pain despite the subjective description of pain. This paper also suggested the use of demographic data in future models, which was done in other studies as aforementioned. This might result in more satisfactory pain control which would improve patient's quality of life [44].

CKD was also investigated by Derebail VK et al. They studied the ability of machine learning to predict eGFR decline, as a marker for kidney function. They also found the best indicators for kidney function decline in SCD patients, which included age, baseline eGFR, and eGFR slope. The predictive ability of the machine learning models was less accurate when trying to predict decline over 12 months when compared to predicting decline over shorter periods like 6 months [45]. Nevertheless, this showed the effectiveness of machine learning to predict a rapid decline of kidney functions, which is defined as a loss of more than 3/mL/min/1.73 mˆ2/ year [46]. This detection might help in risk assessment and modification since this is one of the leading causes of mortality in SCD patients [47]. The ability of AI to predict the occurrence of disease complications serves as a major role in determining individualized disease prognosis and progression. Hence, AI can be used to predict disease prognosis and the specific types of complications that may recur in SCD patients, allowing clinicians to be more aware of potential harm and giving them the ability to predict the disease course more accurately.

2.5. Role of AI in risk stratification

Several studies have also been conducted showing the effect of AI in accurately predicting the risk that SCD patients face based on numerous parameters. Prediction of hospital readmissions and preventing them is an effective way to cut healthcare costs and improve patients' quality of life. Machine-learning algorithms including Logistic Regression, Support-Vector Machine, and Random Forest were used to predict readmissions in patients with SCD. This was compared to the standard readmission scoring systems, including LACE and HOSPITAL indices, and it showed machine-learning superiority over the standard systems [48]. This superiority has been shown over LACE as it can only be used in populations with few comorbidities, which does not match the typical clinical picture of SCD patients. In addition, standard scores usually use a number of predictive features without taking into account disease-specific predictors which may impact the possibility of readmission. For these reasons, readmission scores are not currently routinely used in clinical practice. This might help pinpoint high-risk patients. The study also showed that Random Forest and Logistic Regression algorithms performed the best. This might help in the reallocation of resources to patients who are expected to be readmitted and provide them with sufficient resources to prevent readmission [49]. This helped identify the strongest predictors for readmission [50]. In addition, it was also demonstrated that machine learning methods identified more variables that were associated with SCD patients that were absent in traditional scoring systems for hospital readmissions. For instance, patients living in low-income areas were more susceptible to readmission. Such information will also help adjust any modifiable risks. The predictors used were collected by data-driven methods and clinical knowledge. The most important predictors were also reported. Those included the number of inpatient visits in the past year, the number of days since the last visit, ED visits in the past 6 months, and age. The study also reported those predictors in order according to their importance [49].

Grouping patients in clusters according to their phenotype would provide us with a deeper understanding of the disease pathophysiology and prognosis [51]. The most common cause of hospitalization and morbidity in SCD patients is pain [52]. Using cluster analysis, a study in Brazil was able to categorize patients into five subgroups using ten laboratory tests, showing that hemolytic and inflammatory biomarkers have the greatest influence in the grouping [53]. Clustering algorithms, which is a type of unsupervised machine learning, have also been used to categorize patients into groups according to similarities in clinical data and hence, in prognosis. A study using this technique was able to categorize patients in ICU into three clusters based on disease severity. The grouping of those patients would help in decision-making regarding the risk for each patient group. It would also enhance the coordination in the ICU since it would be more reassuring to transfer less severe patients out of ICU [54].

A study by Sachdev et al. studied the possibility to create a risk score that integrates clinical, laboratory, and imaging results to stratify and assess the risk in patients with SCD. They used data from 600 patients

with SCD and used machine learning techniques to select the best predictors out of 70 variables that were tested. They then performed tests to identify the optimal number of variables to be included. Eventually, they identified nine variables as key independent variables that predict mortality using a multivariable risk model. Those variables were age, body weight, heart rate, left ventricular septal thickness, mitral *E*-wave velocity, tricuspid regurgitant velocity, right atria pressure, blood urea nitrogen, and alkaline phosphatase. These variables showed the effect of the disease on hepatic, renal, and cardiac functions. The study, however, lacked identification of other important disease severity variables like the frequency of vaso-occlusive crises and the need for chronic treatment. The study was also done in a single center and included a relatively small sample size, which reduces its generalizability and external validity. In conclusion, the study was able to provide a risk score to predict mortality that performed well compared to other currently accepted scores [55].

3. Future considerations

In addition, the rise of precision medicine and the increasing popularity of individualizing medicine to each patient instead of according to symptoms signifies the importance of artificial intelligence. It has been shown that artificial intelligence will play an important role as one of the pillars of modern medicine as we move towards precision medicine, especially in genetic diseases like SCD [56,57]. A recent study by El Hoss S et al. has tried to use artificial intelligence in personalized medicine by collecting patient specific pO2 that starts the process of sickling of the red blood cells (PoS), in addition to red blood cell deformability at normoxia (EImax) and during deoxygenation (EImin). They found that there were significant differences in PoS and EImin between patients with the SS genotype and other patients.

Significant differences in all parameters are found between patients on hydroxyurea and patients taking no treatment. A conclusion was reached that these findings can be used to generate AI models that might help in more effective delivery of precision medicine to SCD patients [58]. Currently, it is difficult to use precision medicine in cases of SCD due to the difficulty of classifying patients into groups and the limited availability of treatment [59].

4. Conclusion

Artificial Intelligence use in medicine, and specifically in hemoglobinopathies and SCD, has been on the rise. This review showed the advancements done in diagnostics and risk stratification of sickle cell disease with the help of AI. Moreover, we reviewed the tools being developed to aid clinicians in quantifying and detecting specific disease complications early. We also discuss applications and software that were developed to improve clinical outcomes by enhancing adherence and compliance with treatments. This is directing us towards a new era of faster and more accurate diagnosis, in addition to precise treatment and follow-up according to each SCD patient's specific risks and needs. Through this review, it became clear that AI will inevitably be part of SCD diagnosis and management in clinical practice soon. We recommend further testing of these AI methods on wider populations before they are widely applied. In addition, properly detailed clustering of SCD patients based on clinical and biochemical criteria should be enforced to establish the practicality of using AI. Currently, the main limitations of the application of AI in clinical practice include medicolegal and financial aspects. In addition, generalizability is still in question as more studies on various populations are still required. New fields of medicine, including proteomics, genomics, microbiomes, and metabolomics, are providing large amounts of new information that can be used by artificial intelligence algorithms to further classify patients to provide personalized therapy according to their predicted complications, outcomes, and risk, as this seems to be the future of AI and its applications [59]. Tables 1, 2, and 3 discuss the advantages each gadget or

Table 1
Purpose, advantages, and disadvantages of papers about AI in SCD diagnosis.

Papers	Purpose	Advantages	Limitations
Douglass et al.	Automated sickle cell disease identification using lensless single random phase encoding biosensor (SRPE)	- Low cost - Fast and portable - Can detect diseases other than SCD	- limited by sensor dimensions. - Cannot use 1D crops (not all image types are eligible to be used) - Lack of comparison with existing methods - Lack of diversity in the dataset used
Xu et al.	Uses deep convuluational neural networks for the classification of RBCs in SCD	- showed great accuracy even in deoxygenated datasets - may improve further when the system is better trained. - May be used to identify severity of SCD	- no definite way to automate the inspection and recognition of RBC patterns found in SCD blood samples. - time consuming
Knowlton et al.	smartphone-based sickle cell detection	- fast	- Can only detect HbSS phenotypes - Technical limitations limiting accessibility and applicability
González-Hidalgo M et al.	use of digital data analysis to separate RBC clusters in SCD	- reduces surrounding noise in images. - completely automated. - higher efficiency (100%) compared to other methods.	- Still cannot be used to diagnose. It can only support diagnosis as a method for complementary analysis - Technical limitations: images must meet specific criteria which may not always be feasible in clinical settings.
Cai R et al.	use of active appearance model (AAM) for RBC segmentation	- reduces image noise and separates RBCs from background. - No need to perform boundary tracking when counting cells as the system developed provides a closed form of the boundary. - Incorporates and detects shape and texture, allowing precise segmentation of images of cells.	- Lack of clinical validation
Alagu S et al.	An approach to detect SCD using human RBCs and deep learning with feature selection algorithms.	- Improved Inception V3 model results - Will help in early detection and diagnosis of SCD	- Interpretability: it is hard to understand the method used for the model to make its predictions. - Larger sample size needed to capture all variations and to increase generalizability
Shekhar et al.	use of ACCENSE algorithm and k-means and	- DBSCAN is well studied and widely cited	- Only tried on mice.

(*continued on next page*)

Table 1 (*continued*)

Papers	Purpose	Advantages	Limitations
	DBSCAN to analyze datasets	- Optimizes mass cytometry, which can measure a large number of parameters in individual cells. - Help in automation of classifying subpopulations of cells - Robust algorithm with ability to perform high-dimensional data nalysis	- Requires significant computational resources - designed specifically for transcriptomic data

application provides as well as the drawbacks in each. This is separated into applications in diagnostics, complications diagnosis, and prognosis.

Practice points

- AI has features allowing it to learn, distinguish patterns, and make decisions. It does that by utilizing computational networks like CNN. Machine and deep learning are used in detection of patterns from input data. SCD patients can greatly benefit from the research and algorithms being developed through AI.
- Several AI applications has been developed to aid in SCD diagnosis. This includes usage of a lensless approach called SRPE system to detect and differ sickle cells from healthy cells. dCNNs have also been suggested to classify different shapes or types of RBCs in SCD patient.
- Some methods try to improve accessibility by implementing the utilization of cell phones to aid in diagnosis. Active Appearance Models (AAM) can help in clarifying images and distinguishing cells from background noise.
- AI can be used to analyze retinal imaging and detection of organ dysfunction.
- Medicolegal aspects must be considered when relying on AI for diagnosis.
- Wearable technology is being developed and incorporated to detect vaso-occlusive crises and to predict them. Machine learning can also be used to analyze pain scores during these episodes.
- Machine learning can predict eGFR decline in SCD patients with CKD. It can also help personalize treatments and determine individual risks for patients with SCD.
- Several machine learning algorithms have shown superiority over standard systems in predictions of readmissions.
- Clustering patients according to phenotype provides deeper understanding of the disease pathophysiology and prognosis. This can enhance decision-making and coordination in healthcare.

Research agenda

- Increasing the amount of research using AI in diverse populations in order to enhance generalizability of results.
- Development and refinement of machine learning algorithms to accurately predict prognosis, diagnosis, and complications based on given parameters.
- Exploration of the use of AI and comparing it with established systems and testing its ability to individualize more precise therapies for SCD.
- Investigation of ethical, legal, and social implications of of AI in the context of SCD.

Table 2
Purpose, advantages, and disadvantages of papers about AI in SCD complications.

Cai S et al.	using deep learning to detect sea fan neovascularization from fundus photographs of patients with sickle cell retinopathy	- Study includes a representative population - Fast detection of retinal manifestations - Highly sensitive and specific	- study has single institution design. - small number of images used. - model was trained to detect sea fans only (does not detect another sickle cell retinopathy) - cannot detect stage 4 and 5 sickle cell retinopathies
Mohammed A et al.	predictions are made by machine learning regarding organ dysfunction in ICU admitted SCD patients	- -Early organ failure detection and treatment can reduce mortality rates. - Aids clinical decision making and allows for goal-directed therapy - study addresses a critical issue in the healthcare industry by predicting early-onset acute organ failure in critically ill sickle cell disease patients. The early identification of these patients can significantly improve the treatment and outcomes	- Cannot detect failure of specific organs - Model was built on a small subset of patients - Data used was highly imbalanced (more non-organ failure cases included) - Only patients with at least 24 h of continuous high-frequency physiologic data available before organ failure - Patients who are too sick may not have been connected to monitors - single-center dataset, which may limit the generalizability
Ji Y et al.	identify elevated risk for future pain crises in sickle cell disease patients using photoplethysmogram patterns measured during sleep via a machine learning approach.	- Highly sensitive and specific - Detection and prediction of painful episodes - study uses a non-invasive and convenient method of collecting PPG data during sleep, making it more accessible for patients.	- Patients who are too sick may not have been connected to monitors. - Small dataset - Unbalanced ratios between subjects in each group - It is a single center study, which may limit generalizability.
Hankins et al.	integrating mobile health into sickle cell disease care to increase hydroxyurea utilization	- Has potential to increase patient engagement in hydroxyurea utilization, leading to better outcomes. - Uses data from a diverse population	- study is limited to patients who have access to and are comfortable using mobile health technology - study only focuses on hydroxyurea utilization and does not consider other factors that may affect the patient's overall health outcomes.

(*continued on next page*)

Table 2 (*continued*)

Yang f et al.	using machine learning techniques and physiological measures to improve pain management in patients with sickle cell disease.	- potential to provide a personalized and objective pain management strategy for patients and improve pain management - may lead to development of non-invasive and continuous pain monitoring system.	- Small sample size - physiological measures used in the study were limited to heart rate variability and electrodermal activity, and other relevant measures may need to be incorporated for better accuracy.
Alambu et al.	using natural language processing to measure pain in SCD patients using clinical text.	- non-invasive and efficient way of collecting data on pain experiences and management in clinical settings. - Efficient: may save time and resources - May improve pain management	- non-invasive and efficient way of collecting data on pain experiences and management in clinical settings. - It is a single center study, which may limit generalizability.
Johnson et al.	using mobile health apps and wearable technology to assess changes and predict pain during treatment of acute pain in SCD	- includes a range of outcome measures, including pain severity, physical activity, and sleep, providing a comprehensive assessment of the intervention's effectiveness. - potential to provide a personalized and objective pain management strategy for patients and improve pain management.	- study is limited to patients who have access to and are comfortable using mobile health apps and wearable technology - study relies on patient self-reporting of pain severity, which may be subjective and influenced by various factors
Sachdev et al.	Creating a risk score using machine learning to predict mortality	- outperformed currently accepted risk scores. - Identifies the most important predictable factors that may predict mortality. - Allows lower cost methods for institutioms with less resources	- Low external validity due to small sample size and being a single center study. - Did not include important factors that shows disease severity such as pain crises and chronic treatments like hydroxyurea and transfusions.

Contribution to the field statement

The authors created a comprehensive review to guide clinicians and hematologists. This is one of the first reviews to discuss, compare, and contrast different novel methods that utilize AI in the diagnosis and treatment of SCD. We also discuss the future implications of AI in the field of hematology and specifically in SCD. We believe that this paper would help clinicians and researchers in decision making and critical appraisal of all the published data on AI methods in SCD and would aid in the appraisal of future methods, too.

Table 3
Purpose, advantages, and disadvantages of papers about AI in SCD Risk Stratification.

Patel et al.	utilizes machine learning algorithms to predict hospital re-admissions in SCD patients	- The approach has the potential to improve patient outcomes and reduce healthcare costs. - study uses a large dataset	study does not account for social determinants of health, which may influence hospital re-admissions. accuracy of the machine learning algorithms may be affected by missing data and incomplete medical records.
Padrão EMH et al.	uses an unsupervised machine learning approach to identify phenotypes of sickle cell intensive care admissions.	- improve understanding of sickle cell disease progression and guide personalized treatment strategies. - study uses a large dataset	study does not account for social determinants of health, which may influence hospital re-admissions. accuracy of the machine learning algorithms may be affected by missing data and incomplete medical records.
Dutra et al.	utilizes hierarchical cluster analysis to identify clinical profiles in a cohort of sickle cell anemia patients in Brazil	- The approach may help identify patient subgroups and guide personalized treatment strategies	sample size is relatively small

Author contributions

The Conceptualization, A.A.E, M.Y; Writing - original draft, A.A.E, M.Y.; methodology, A.A.E, A.M.E M.E. and B.E.; software, A.A.E.; validation, R.K., A.A., A.A.A., and M.A.; investigation, M.E., A.A.E. and K. F..; resources, B.E and A.M.E.; writing—review and editing, A.E., A.M. E., B.E., and A.A.; supervision, M.Y., S.A., J.A.A; project administration, A.A.E., M.Y., R.K.,K.F. All authors have read and agreed to the published version of the manuscript.

Funding

Open Access funding provided by the Qatar National Library.

Declaration of Competing Interest

All The authors declare no conflict of interest.

References

[1] Chen ZH, Lin L, Wu CF, Li CF, Xu RH, Sun Y. Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine. Cancer Commun (Lond) 2021;41(11):1100–15.
[2] Muthalaly RG, Evans RM. Applications of machine learning in cardiac electrophysiology. Arrhythmia Electrophysiol Rev 2020;9(2):71–7.
[3] Majumder A, Sen D. Artificial intelligence in cancer diagnostics and therapy: current perspectives. Indian J Cancer 2021;58(4):481–92.
[4] Elsayed B, Elshoeibi AM, Elhadary M, Ferih K, Elsabagh AA, Rahhal A, et al. Applications of artificial intelligence in Philadelphia-negative myeloproliferative neoplasms. Diagnostics. 2023;13(6):1123.
[5] Elshoeibi AM, Ferih K, Elsabagh AA, Elsayed B, Elhadary M, Marashi M, et al. Applications of artificial intelligence in thrombocytopenia. Diagnostics. 2023;13(6):1060.

[6] Busnatu Ș, Niculescu AG, Bolocan A, Petrescu GED, Păduraru DN, Năstasă I, et al. Clinical applications of artificial intelligence-an updated overview. J Clin Med 2022;11(8).
[7] Ata F, Yousaf Z, Sardar S, Javed S, Iqbal P, Khamees I, et al. Protocol for "genetic composition of sickle cell disease in the Arab population: a systematic review". Health Sci Rep 2022;5(3):e450.
[8] Inusa BPD, Hsu LL, Kohli N, Patel A, Ominu-Evbota K, Anie KA, et al. Sickle cell disease-genetics, pathophysiology, clinical presentation and treatment. Int J Neonatal Screen 2019;5(2):20.
[9] Das PK, Meher S, Panda R, Abraham A. A review of automated methods for the detection of sickle cell disease. IEEE Rev Biomed Eng 2020;13:309–24.
[10] Ismail A, Yusuf AA, Kuliya-Gwarzo A, Ahmed SG, Tabari AM, Abubakar SA. Correlating transcranial arterial Doppler velocities with haematologic parameters and haemolytic indices of Nigerian children with sickle cell anaemia. Ultrasound. 2019;27(2):101–10.
[11] David AN, Jinadu MY, Wapmuk AE, Gbajabiamila TA, Okwuzu JO, Herbertson EC, et al. Prevalence and impact of sickle cell trait on the clinical and laboratory parameters of HIV infected children in Lagos. Nigeria Pan Afr Med J 2018;31:113.
[12] Yassin M, Soliman A, De Sanctis V, Nashwan A, Abusamaan S, Moustafa A, et al. Liver iron content (LIC) in adults with sickle cell disease (SCD): correlation with serum ferritin and liver enzymes concentrations in Trasfusion dependent (TD-SCD) and non-transfusion dependent (NT-SCD) patients. Mediterr J Hematol Infect Dis 2017;9(1):e2017037.
[13] Rozi W, Rahhal A, Ali EA, Al-Mashdali A, Hilan Y, Khamees I, et al. Direct oral anticoagulants in sickle cell disease: a systematic review and meta-analysis. Blood Adv 2022;6(17):5061–6.
[14] Chakravorty S, Williams TN. Sickle cell disease: a neglected chronic disease of increasing global health importance. Arch Dis Child 2015;100(1):48–53.
[15] Shier A, Abdelrazek M, Soliman A, De Sanctis V, Elsayed A, Abdulla M, et al. Short-term outcome and MRI changes in three adult patients with sickle cell disease and aseptic osteonecrosis after treatment with hyperbaric oxygen therapy: a preliminary report. Case Rep Oncol 2020;13(1):365–72.
[16] Alshurafa A, Yassin MA. Case report: safety and efficacy of voxelotor in a patient with sickle cell disease and stage IV chronic kidney disease. Front Med (Lausanne) 2022;9:931924.
[17] Mazandu GK, Hotchkiss J, Nembaware V, Wonkam A, Mulder N. The sickle cell disease ontology: recent development and expansion of the universal sickle cell knowledge representation. Database (Oxford) 2022;2022.
[18] Douglass PM, O'Connor T, Javidi B. Automated sickle cell disease identification in human red blood cells using a lensless single random phase encoding biosensor and convolutional neural networks. Opt Express 2022;30(20):35965–77.
[19] Knowlton SM, Sencan I, Aytar Y, Khoory J, Heeney MM, Ghiran IC, et al. Sickle cell detection using a smartphone. Sci Rep 2015;5:15022.
[20] Xu M, Papageorgiou DP, Abidi SZ, Dao M, Zhao H, Karniadakis GE. A deep convolutional neural network for classification of red blood cells in sickle cell anemia. PLoS Comput Biol 2017;13(10):e1005746.
[21] González-Hidalgo M, Guerrero-Peña FA, Herold-García S, Jaume-I-Capó A, Marrero-Fernández PD. Red blood cell cluster separation from digital images for use in sickle cell disease. IEEE J Biomed Health Inform 2015;19(4):1514–25.
[22] Cai R, Wu Q, Zhang R, Fan L, Ruan C, editors.. Red blood cell segmentation using active appearance model. 2012 IEEE 11th international conference on signal processing. 2012. p. 21–5. Oct. 2012.
[23] Darrin M, Samudre A, Sahun M, Atwell S, Badens C, Charrier A, et al. Classification of red cell dynamics with convolutional and recurrent neural networks: a sickle cell disease case study. Sci Rep 2023;13(1):745.
[24] Shekhar K, Brodin P, Davis MM, Chakraborty AK. Automatic classification of cellular expression by nonlinear stochastic embedding (ACCENSE). Proc Natl Acad Sci U S A 2014;111(1):202–7.
[25] Li H, Zheng W-S, Zhang J. Deep CNNs for HEp-2 cells classification: a cross-specimen analysis. arXiv preprint arXiv:160405816. 2016.
[26] Ganesan K. A novel deep learning approach for sickle cell anemia detection in human RBCs using an improved wrapper-based feature selection technique in microscopic blood smear images. Biomed Tech (Berl) 2022;68(1).
[27] Cai S, Parker F, Urias MG, Goldberg MF, Hager GD, Scott AW. Deep learning detection of sea Fan neovascularization from ultra-widefield color fundus photographs of patients with sickle cell hemoglobinopathy. JAMA Ophthalmol 2021;139(2):206–13.
[28] Scruggs BA, Chan RVP, Kalpathy-Cramer J, Chiang MF, Campbell JP. Artificial intelligence in retinopathy of prematurity diagnosis. Transl Vis Sci Technol 2020;9(2):5.
[29] Sevgi DD, Scott AW, Martin A, Mugnaini C, Patel S, Linz MO, et al. Longitudinal assessment of quantitative ultra-widefield ischaemic and vascular parameters in sickle cell retinopathy. Br J Ophthalmol 2022;106(2):251–5.
[30] Wisely CE, Wang D, Henao R, Grewal DS, Thompson AC, Robbins CB, et al. Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging. Br J Ophthalmol 2022;106(3):388–95.
[31] Condon PI, Serjeant GR. Behaviour of untreated proliferative sickle retinopathy. Br J Ophthalmol 1980;64(6):404–11.
[32] Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019;380(14):1347–58.
[33] Mohammed A, Podila PSB, Davis RL, Ataga KI, Hankins JS, Kamaleswaran R. Using machine learning to predict early onset acute organ failure in critically ill intensive care unit patients with sickle cell disease: retrospective study. J Med Internet Res 2020;22(5):e14693.
[34] Hankins JS, Shah N, DiMartino L, Brambilla D, Fernandez ME, Gibson RW, et al. Integration of Mobile health into sickle cell disease care to increase hydroxyurea utilization: protocol for an efficacy and implementation study. JMIR Res Protoc 2020;9(7):e16319.
[35] Kamaleswaran R, Akbilgic O, Hallman MA, West AN, Davis RL, Shah SH. Applying artificial intelligence to identify Physiomarkers predicting severe Sepsis in the PICU. Pediatr Crit Care Med 2018;19(10):e495–503.
[36] Ji Y, Chalacheva P, Rosen CL, DeBaun MR, Coates TD, Khoo MCK. Identifying elevated risk for future pain crises in sickle-cell disease using photoplethysmogram patterns measured during sleep: a machine learning approach. Front Digit Health 2021:3.
[37] Johnson A, Yang F, Gollarahalli S, Banerjee T, Abrams D, Jonassaint J, et al. Use of mobile health apps and wearable technology to assess changes and predict pain during treatment of acute pain in sickle cell disease: feasibility study. JMIR Mhealth Uhealth 2019;7(12):e13671.
[38] Vatansever S, Schlessinger A, Wacker D, Kaniskan H, Jin J, Zhou MM, et al. Artificial intelligence and machine learning-aided drug discovery in central nervous system diseases: state-of-the-arts and future directions. Med Res Rev 2021;41(3):1427–73.
[39] Issom DZ, Rochat J, Hartvigsen G, Lovis C. Preliminary evaluation of a mHealth coaching conversational artificial intelligence for the self-care management of people with sickle-cell disease. Stud Health Technol Inform 2020;270:1361–2.
[40] Badawy SM, Cronin RM, Hankins J, Crosby L, DeBaun M, Thompson AA, et al. Patient-centered eHealth interventions for children, adolescents, and adults with sickle cell disease: systematic review. J Med Internet Res 2018;20(7):e10940.
[41] Green NS, Manwani D, Matos S, Hicks A, Soto L, Castillo Y, et al. Randomized feasibility trial to improve hydroxyurea adherence in youth ages 10-18 years through community health workers: the HABIT study. Pediatr Blood Cancer 2017;64(12).
[42] Smaldone A, Findley S, Manwani D, Jia H, Green NS. HABIT, a randomized feasibility trial to increase hydroxyurea adherence, suggests improved health-related quality of life in youths with sickle cell disease. J Pediatr 2018;197:177–85.e2.
[43] Alambo A, Andrew R, Gollarahalli S, Vaughn J, Banerjee T, Thirunarayan K, et al. Measuring pain in sickle cell disease using clinical text. Annu Int Conf IEEE Eng Med Biol Soc 2020;2020:5838–41.
[44] Yang F, Banerjee T, Narine K, Shah N. Improving pain management in patients with sickle cell disease from physiological measures using machine learning techniques. Smart Health (Amst) 2018;7-8:48–59.
[45] Güntürkün F, Chen D, Akbilgic O, Davis RL, Karabayir I, Strome M, et al. Using machine learning to predict rapid decline of kidney function in sickle cell anemia. EJHaem. 2021;2(2):257–60.
[46] Derebail VK, Ciccone EJ, Zhou Q, Kilgore RR, Cai J, Ataga KI. Progressive decline in estimated GFR in patients with sickle cell disease: an observational cohort study. Am J Kidney Dis 2019;74(1):47–55.
[47] Shlipak MG, Katz R, Kestenbaum B, Siscovick D, Fried L, Newman A, et al. Rapid decline of kidney function increases cardiovascular risk in the elderly. J Am Soc Nephrol 2009;20(12):2625–30.
[48] Brennan JJ, Chan TC, Killeen JP, Castillo EM. Inpatient readmissions and emergency department visits within 30 days of a hospital admission. West J Emerg Med 2015;16(7):1025–9.
[49] Patel A, Gan K, Li AA, Weiss J, Nouraie M, Tayur S, et al. Machine-learning algorithms for predicting hospital re-admissions in sickle cell disease. Br J Haematol 2021;192(1):158–70.
[50] Chai SY, Hayat A, Flaherty GT. Integrating artificial intelligence into haematology training and practice: opportunities, threats and proposed solutions. Br J Haematol 2022;198(5):807–11.
[51] Coelho A, Dias A, Morais A, Nunes B, Faustino P, Lavinha J. Sickle cell disease severity scoring: a yet unsolved problem. Eur J Haematol 2012;89(6):501–2.
[52] Sant'Ana PG, Araujo AM, Pimenta CT, Bezerra ML, Junior SP, Neto VM, et al. Clinical and laboratory profile of patients with sickle cell anemia. Rev Bras Hematol Hemoter 2017;39(1):40–5.
[53] Dutra VF, Biassi TP, Figueiredo MS. Sickle cell anemia: hierarchical cluster analysis and clinical profile in a cohort in Brazil. Hematol Transfus Cell Ther 2021 2023;68(2):175–85.
[54] Padrão EMH, Bustos B, Mahesh A, Fonseca GHH, Taniguchi LU. Phenotypes of sickle cell intensive care admissions: an unsupervised machine learning approach in a single-center retrospective cohort. Ann Hematol 2022;101(9):1951–7.
[55] Sachdev V, Tian X, Gu Y, Nichols J, Sidenko S, Li W, et al. A phenotypic risk score for predicting mortality in sickle cell disease. Br J Haematol 2021;192(5):932–41.
[56] Abdelhalim H, Berber A, Lodi M, Jain R, Nair A, Pappu A, et al. Artificial intelligence, healthcare, clinical genomics, and pharmacogenomics approaches in precision medicine. Front Genet 2022;13:929736.
[57] Alrefaei AF, Hawsawi YM, Almaleki D, Alafif T, Alzahrani FA, Bakhrebah MA. Genetic data sharing and artificial intelligence in the era of personalized medicine based on a cross-sectional analysis of the Saudi human genome program. Sci Rep 2022;12(1):1405.
[58] A I, A CG, E V, R C, P B, M DEM, et al. P-032: Integrative diagnosis of sickle cell disease patients for personalized medicine. 2023. LID - 33. (2572–9241 (Electronic)).
[59] El Hoss S, El Nemer W, Rees DC. Precision medicine and sickle cell disease. Hemasphere. 2022;6(9):e762.