SYSTEMATIC REVIEW

DOI: http://dx.doi.org/10.15403/jgld-4755

# Applications of Artificial Intelligence in the Automatic Diagnosis of Focal Liver Lesions: A Systematic Review

**Stefan Lucian Popa[1], Simona Grad[1], Giuseppe Chiarioni[2], Annalisa Masier[3], Giulia Peserico[3], Vlad Dumitru Brata[4], Dinu Iuliu Dumitrascu[5], Alberto Fantin[3]**

1) 2nd Department of Internal Medicine, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania;
2) Division of Gastroenterology B, AOUI Verona, Verona, Italy and Division of Gastroenterology and Hepatology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA;
3) Division of Gastroenterology, Department of Oncology, Veneto Institute of Oncology IOV – IRCCS, Padua, Italy;
4) Faculty of Medicine, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania;
5) Department of Anatomy, UMF Iuliu Hatieganu Cluj-Napoca, Cluj-Napoca, Romania.

***Address for correspondence:***
Simona Grad
2nd Department of Internal Medicine, „Iuliu Hatieganu" University of Medicine and Pharmacy,
Cluj-Napoca, Romania;
costinsimona_m@yahoo.com

Received: 01.12.2022
Accepted: 15.02.2023

## ABSTRACT

**Background & Aims**: Focal liver lesions (FLLs) are defined as abnormal solid or liquid masses differentiated from normal liver, frequently being clinically asymptomatic. The aim of this systematic review is to provide a comprehensive overview of current artificial intelligence (AI) applications, deep learning systems and convolutional neural networks, capable of performing a completely automated diagnosis of FLLs.
**Methods**: We searched PubMed, Cochrane Library, EMBASE, and WILEY databases using predefined keywords. Articles were screened for relevant publications about AI applications capable of automated diagnosis of FLLs. The search terms included: (focal liver lesions OR FLLs OR hepatic focal lesions OR liver focal lesions OR liver tumor OR hepatic tumor) AND (artificial intelligence OR machine learning OR neural networks OR deep learning OR automated diagnosis OR ultrasound OR US OR computer scan OR CT OR magnetic resonance imaging OR MRI OR computer-aided diagnosis OR automated computer tomography OR automated magnetic imaging).
**Results**: Our search identified a total of 32 articles analyzing complete automated imagistic diagnosis of FLLs, out of which 14 studies analyzing liver ultrasound images, 8 studies analyzing computer tomography images and 10 studies analyzing images obtained from magnetic resonance imaging.
**Conclusions**: We found significant evidence demonstrating that implementing a complete automated system for FLLs diagnosis using AI-based applications is currently feasible. Various automated AI-based applications have been analyzed. However, there is no clear evidence about the superiority of any of the systems.

**Key words: f**ocal liver lesions – artificial intelligence – hepatic tumors – focal hepatic lesions – machine learning – neural networks – deep learning – automated diagnosis – computer-aided diagnosis.

**Abbreviations**: AI: artificial intelligence; AUC: area under the curve; B-US: B-mode ultrasound; CAD: computer-aided diagnostic; CT: computer tomography; CEUS: contrast enhanced ultrasound; DIA: digital image analysis; DL: Deep Learning; ELM extreme learning machine; EUS: endoscopic ultrasound; FLL: focal liver lesion; FNH: focal nodullar hyperplasia; HCC: hepatocellular carcinoma; ICC: intrahepatic cholangiocarcinoma; IQR interquartile range; MRI: magnetic resonance imaging; MFI: microflow imaging; SVM: support vector machine; WSI: whole slide image.

## INTRODUCTION

Focal liver lesions (FLLs) are defined as abnormal solid or liquid masses differentiated from the normal liver. Frequently, FFLs are clinically asymptomatic and are detected incidentally by imaging due to unrelated symptoms, using the following imaging techniques: B-mode ultrasound, contrast enhanced ultrasound (CEUS), elastography, contrast enhanced computed tomography (CT) scan and contrast enhanced magnetic resonance imaging (MRI) [1]. Nevertheless, large FFLs may be associated with right upper quadrant abdominal pain or bleeding.

Epidemiologic studies indicate that FLLs were found in up to 33% of radiological studies and in more than 50% autopsy cases [2].

Focal liver lesions can be categorized as solid, which are typically benign or malignant tumors, liquid (cystic), which are typically liquid filled cavities or abscesses, and hemangioma, which may have imaging characteristics of either cystic or solid lesions. Benign lesions are classified as hepatic hemangioma, focal nodular hyperplasia (FNH), hepatocellular adenoma, regenerative nodules, and malignant lesions as hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (ICC), metastatic disease or rare liver cancers [3].

The differential diagnosis is a complex process and requires the understanding of the clinical context with a detailed history, physical examination, blood and urine tests and finally choosing the correct imaging technique. If the FLLs are found incidentally, the first step in the management is to exclude HCC, as well as liver cirrhosis, chronic hepatitis, autoimmune hepatitis and B or C viral infection [3-5]. Further, imaging-guided liver biopsy is still recommended, if the diagnosis is uncertain after obtaining contrast-enhanced, cross-sectional MRI imaging [1-5]. Notwithstanding the high diagnosis precision of liver biopsy, it is rarely recommended due to an increased risk of life-threatening complications.

Unfortunately, there is no clear pathway for a work up, and, with a wide differential diagnosis, these lesions may need multiple imaging modalities to characterize whether they are benign or malignant.

Because of the absence of a precise imagistic method of diagnosis, there is no consensus of a clear pathway for work up and very often the available techniques are inaccurate and highly expensive. For this reason, automated diagnosis systems based on artificial intelligence (AI) applications were developed, to provide a precise diagnosis in a time saving, cost efficient future perspective.

This aim of this systematic review is to provide a comprehensive overview of current AI applications, deep learning systems and convolutional networks, which can perform a completely automated diagnosis of FLLs.

## METHODS

We searched PubMed, Cochrane Library, EMBASE, and WILEY databases using predefined keywords. Articles were screened for relevant publications about AI applications capable of automated diagnosis of FLLs. The search terms included: (focal liver lesions OR FLLs OR hepatic focal lesions OR liver focal lesions OR liver tumor OR hepatic tumor) AND (artificial intelligence OR machine learning OR neural networks OR deep learning OR automated diagnosis OR ultrasound OR US OR computer scan OR CT OR magnetic resonance imaging OR MRI OR computer-aided diagnosis OR automated computer tomography OR automated magnetic imaging). Exclusion criteria were case reports, pediatric studies, abstracts, conference presentations, letters to the editor, studies written in languages other than English, and editorials.

Four independent authors (S.L.P., S.G., D.I.D., and V.D.B.) reviewed for eligibility titles, abstracts, full text of eligible articles. Data extraction was conducted independently by all four reviewers. Any discrepancies in extracted data were resolved by mutual consensus. Extracted data on the authors' names, year of publication, country or study population, sample size, study design, the method used to diagnose FLLs, artificial intelligence-based application were reported into three separate tables. Fig. 1 shows the search strategy using the PRISMA flow diagram.

## RESULTS

Our search identified a total of 32 studies analyzing complete automated imagistic diagnosis of FLLs, out of which 14 studies analyzing liver ultrasound images, 8 studies

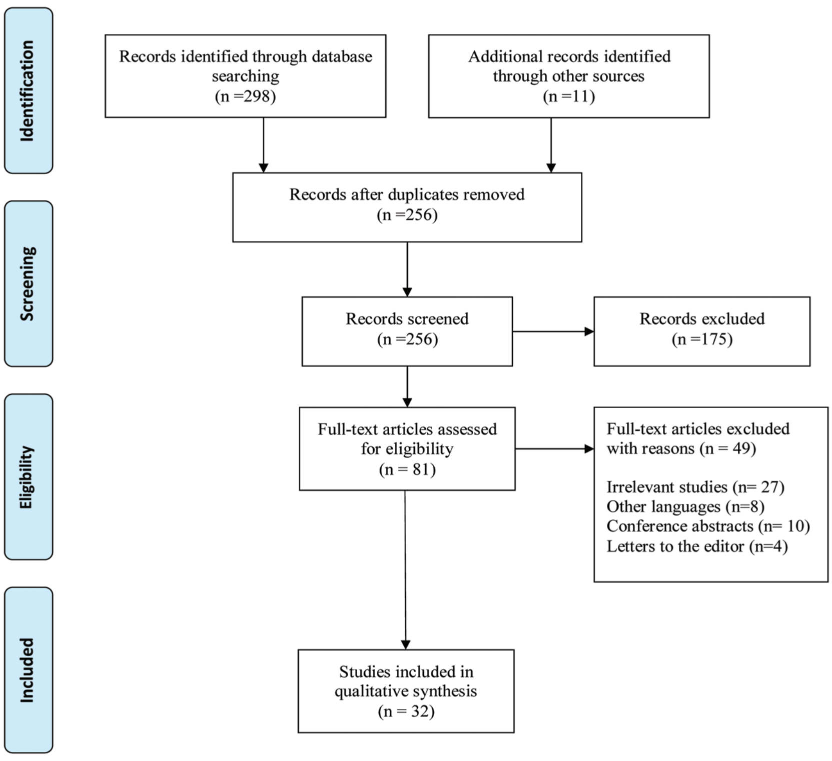


**Fig. 1**. PRISMA flow diagram

analyzing CT images and 10 studies analyzing images obtained from MRI, as demonstrated in Fig. 1.

**Artificial Intelligence based on Liver Ultrasound**

A study performed by Gatos et al. [6] analyzed an automated quantification algorithm for the detection and classification of FLLs using CEUS [6]. 52 CEUS video sequences (30 benign and 22 malignant) were included in the study and the algorithm was programmed to detect contour focal lesions [6]. The results showed a value of 90.3% for classification accuracy with sensitivity and specificity values of 93.1% and 86.9%, respectively [6]. Three benign and two malignant FLLs were misdiagnosed [6]. All studies analyzing AI-based systems using liver ultrasound images for an automated diagnosis of FLLs are in Table I.

Hwang et al. [7] analyzed the possibility of classifying FLLs by extracting hybrid textural features and using an artificial neural network from ultrasound images. A total of 99 images of FLLs were included in the study. Focal liver lesions were divided into 29 cysts, 37 hemangiomas, and 33 malignancies and 29 key features that were selected as a set of inputs for the artificial neural network [7]. Positive predictive value, negative predictive value, sensitivity, specificity, and accuracy were analyzed for each FLL [7]. The results indicate that the algorithm exhibits a high diagnosis accuracy of over 96% among all FLLs groups and the accuracy was increased when echogenicity was included in the feature set [7]. The authors concluded that their system can currently be applied clinically, considering the high diagnosis accuracy.

An endoscopic ultrasound (EUS)-based convolutional neural network was used to distinguish between benign and malignant hepatic masses in a study performed by Marya et al. [8], which analyzed a prospective EUS database with a total of 210,685 images from 256 patients. The data was used to test, train, and further validate the autonomous system. The results showed that the convolutional neural network was successful in autonomously locating FLLs in 92% of EUS video when occlusion heatmap analyzing option was added [8].

**Table I**. Studies assessing automated diagnosis of FLLs using liver ultrasound images

| Study | Publication Year | Total number of images/patients | Diagnosis | Main Findings |
|---|---|---|---|---|
| Gatos et al. [6] | 2015 | 52 images | Benign and malignant FLLs. | Highest classification accuracy for FLLs from the support vector machine model was 90.3%. Five FLLs were misdiagnosed. |
| Hwang et al. [7] | 2015 | 99 images | Cysts, hemangiomas, malignancies | The method shows a high diagnosis accuracy: 96% among all FLLs |
| Marya et al. [8] | 2020 | 210,685 images | Benign and malignant FLLs | The method was 100% sensitive and 80% specific for classifying malignant FLLs, if full length videos were used. |
| Mittal et al. [9] | 2011 | 111 images | Cyst, hemangioma, HCC, metastases, normal liver | The system reported a correct diagnosis in 90.3% of the cases and the results of two-step neural network classifier showed correct decisions of 432 out of 500 segmented regions-of-interest. |
| Virmani et al. [10] | 2014 | 108 images | Cyst, hemangioma, metastatic carcinoma lesions, small and large HCC, normal liver | The first step of the classification module accuracy was 88.7 % and the second step was 95%. |
| Streba et al. [11] | 2012 | 112 images | HCC, liver metastases, hepatic hemangiomas, focal fatty changes | The neural network had 94.45% training accuracy and 87.12% testing accuracy, providing a reliable diagnostic tool for the clinician. |
| Sugimoto et al. [12] | 2009 | 137 images | HCC (23: well-differentiated, 36: moderately differentiated, 15: poorly differentiated), liver metastases, liver hemangiomas | The automated system shows an accuracy of 84.8% for metastasis, 93.3% for hemangioma, and 98.6% for all HCCs, demonstrating the efficiency of a computer-aided diagnostic scheme. |
| Shiraishi et al. [13] | 2008 | 103 images | Liver metastases, HCC | The system showed an accuracy of 88.5% for metastasis, 93.8% for hemangioma, and 86.9% for all hepatocellular carcinomas. |
| Schmauch et al. [14] | 2019 | 367 images | Benign and malignant FLLs | The models reached mean ROC-AUC scores of 0.935 for FLL detection and 0.916 for FLL characterization. |
| Acharya et al. [15] | 2018 | 140 images | Benign and malignant FLLs | Compared with other computer-aided diagnostic–based systems, the method is fully autonomous, because no segmentation of region-of-interest is needed. |
| Xi et al. [16] | 2021 | 911 images | Benign and malignant FLLs | The test accuracy of the model was 0.84, while the accuracy on the uncertain set of images was 0.79, outperforming the other two experts in both situations. |
| Ryu et al. [17] | 2021 | 4,309 images | Hepatic cysts, hemangioma, liver metastasis, HCC | The model achieved a sensitivity of 86.7%, specificity of 89.7%, and an AUC of 0.947. |
| Tiyarattanachai et al. [18] | 2021 | 40,397 images | HCC, cysts, hemangiomas, focal fatty sparing, focal fatty infiltration | The sensitivity, specificity and overall detection rate were 83.9%, 97.1%, 87% for the internal test, and 84.9%, 97.1%, and 75% for the external validation. |
| Mao et al. [19] | 2021 | 114 patients | HCC, ICC, liver metastases | One of the five algorithms achieved a sensitivity of 0.768 and a specificity of 0.88 when differentiating between primary and metastatic liver cancer. |

AUC: area under curve; FLL: focal liver lesion; HCC: hepatocellular carcinoma; ICC: intrahepatic cholangiocarcinoma; ROC: receiver operating characteristic.;

Another study analyzing a system based on neural networks using US images for the autonomous diagnosis of FLLs was conducted by Mittal et al. [9] and included 88 subjects and a total of 111 ultrasound images. The system was programmed to diagnose the following types of FLLs: cyst, hemangioma, HCC, metastases, and normal hepatic images. The results showed correct diagnosis in 432 out of 500 and a classification accuracy of 86.4%, out of which 90.3% (308/340) in typical cases and 77.5% (124/160) in atypical cases [9]. A neural network ensemble-based computer-aided diagnostic (CAD) system used for the differential diagnosis of FLLs was studied by Virmani et al. [10] and the result was that diagnosis accuracy increases from 88.7% to 95% if the second step of the classification module was included in the analysis.

Streba et al. [11] analyzed CEUS imaging, focusing on the role of time-intensity curve parameters involved in a system of neural networks programmed for automatic diagnosis of FLLs. Full length videos of contrast uptake phases were recorded and were analyzed by the neural network. Further, a ratio between median intensities of the central and peripheral areas was autonomously analyzed by a multi-layer neural network which classified the videos into five types, corresponding to each type of liver lesion. The results showed that the neural network had a training accuracy of 94.45% and a testing accuracy of 87.12%. The automatic classification process showed a 93.2% sensitivity, 89.7% specificity, 94.42% positive predictive value and 87.57% negative predictive value. The authors concluded that the accuracy of the autonomous system was similar to that of human interpretation of the time-intensity curves ($p=0.225$ and $p=0.451$, respectively) [11].

Sugimoto et al. [12] used four different artificial neural networks in order to develop a CAD system for automatic diagnosis of FLLs using hepatic CEUS images [12]. In the study 137 patients were included and classification accuracies were 84.8% for metastasis, 93.3% for hemangioma, and 98.6% for all HCCs.

A system based on microflow imaging of CEUS was applied by Shiraishi et al. [12] on 97 images of FLLs. Further, six independent artificial neural networks were involved in process of autonomous diagnosis. The results showed an accuracy of 88.5% for metastasis, 93.8% for hemangioma, and 86.9% for all HCCs [13].

Deep learning was utilized for a system of diagnosis of FLLs in a study performed by Schmauch et al. [13], using 367 two-dimensional hepatic ultrasound images and results showed a mean area under the receiver operating characteristic (AUROC) curve scores of 0.935 for FLL detection and 0.916 for FLL characterization over three shuffled three-fold cross-validations performed with the training data [14].

Acharya et al. [15] analyzed the feasibility of radon transform and bi-directional empirical mode decomposition diagnose FLLs, and the extracted parameters were included in a process of particle swarm optimization for the extraction of optimized details. The system was trained using 78 normal liver images, 26 benign FLLs and 36 malignant FLLs [15]. The results showed an accuracy of 92.95%, a sensitivity of 90.80%, and a specificity of 97.44% [15]. Compared with other AI-based systems, this CAD system is fully automatic because no segmentation of region-of-interest is needed.

Xi et al. [16] developed a model based on machine learning in order to differentiate between benign and malignant liver lesions on ultrasonography. To this extent, the study was conducted on 596 patients, 376 benign and 535 malignant images respectively. Moreover, the training set consisted of 660 lesions, and the test set contained 79 images. Overall, the model outperformed the human experts in analyzing every set of images, with a test accuracy of 0.84 (compared with 0.8 and 0.73), and an accuracy in analyzing the uncertain set of lesions of 0.79 (compared with 0.7 and 0.66) [16].

Ryu et al. [17] developed an algorithm based on convolutional neural networks in order to differentiate between several FLLs, such as hepatic cysts, hemangiomas, liver metastasis and HCC [17]. Overall, the study was based on 4,309 images from 3,873 patients confirmed with those pathologies, among which 3,909 were used for training, and 400 for testing respectively. Thus, the program achieved a sensitivity of 86.7%, specificity of 89.7%, and an AUC of 0.947 [17].

In comparison, Tiyarattanachai et al. [18] also used convolutional neural networks in order to detect and diagnose FLLs, using 40,397 images from 3,487 patients in order to train the model, and an additional 6191 images for an internal test and 18,922 images for an external test respectively [18]. The sensitivity, specificity and overall detection rate were 83.9%, 97.1%, 87% for the internal test, and 84.9%, 97.1%, and 75% for the external validation as well. It is noteworthy that the algorithm performed very well when detecting HCC, achieving a sensitivity of 81.5%, a specificity of 94.4%, and a negative predictive value of 97.4% for the external set [18].

Mao et al. [19] used machine learning-based radiomics in order to create several algorithms that would differentiate between primary and metastatic liver cancer. Thus, five models using random forest, k-nearest neighbours, multilayer perception, support vector machine, and logistic regression respectively were built. Out of these models, the logistic regression managed to differentiate between primary and metastatic liver cancer with a sensibility of 0.816 and a specificity of 0.768 [19].

All studies analyzing AI-based systems using liver ultrasound images for an automated diagnosis of FLLs are in Table I.

**Artificial Intelligence based on Computer Tomography**

We found 8 articles analyzing automated diagnosis using CT images (Table II).

A study performed by Massoptier et al. [20] analyzed an autonomous diagnosis system using CT scans. The method does not require interaction between the physician, nurse or imaging technician and the AI-based system because the initialization of the algorithm is fully automatic. A statistical model-based application was added to detect the liver and to make a precise difference from other abdominal organs. Further, an active contour technique with gradient vector flow was additionally included in the software of the system to obtain a precise hepatic segmentation [20]. The results showed a sensitivity of 82.6% and specificity of 87.5% for tumor lesion detection and the system had a high accuracy of diagnosis with a short processing time (11.4 s for a 512 x 512-pixel slice) [20].

**Table II**. Studies assessing automated diagnosis of focal liver lesions using computer tomography scan images

| Study | Publication Year | Total number of images/patients | Diagnosis | Main Findings |
|---|---|---|---|---|
| Massoptier et al. [20] | 2008 | 21 images | Hepatic tumors | Additionally, the system combined an active contour technique using gradient vector flow in order to obtain a more precise liver surface segmentation. The results showed a sensitivity for tumor lesion detection of 82.6% and a specificity 87.5%, respectively. |
| Bilello et al. [21] | 2004 | 56 images | Simple cysts, hemangiomas, metastases, cyst | The method showed good results for the automated diagnosis of cysts, hemangioma, but was least precise in performing a differential diagnosis between hemangiomas and metastases. |
| Yang et al. [22] | 2012 | 189 images | Hepatomas, cysts, hemangiomas | The system used a content-based retrieval method using bag-of-visual-words representations of single and multiple phases of images obtained from computer imaging scans. The results showed the mean of average precision reach more than 90%. |
| Shi et al. [23] | 2020 | 449 images | HCC and non-HCC groups | The diagnostic accuracy in differentiating HCC from other FLLs on test sets was 83.3% for method A (four-phase CT images) A, 81.1% for method B (three-phase images without portal-venous phase) and 85.6% for method C (three-phase images without precontrast phase). |
| Gao et al. [24] | 2021 | 723 patients | HCC, ICC, metastatic liver cancer | On the test set aimed at identifying hepatocellular carcinoma and intrahepatic cholangiocarcinoma, the deep learning model achieved an accuracy of 86.2%, with an AUC of 0.893. When it comes to the differential diagnosis between the two types of cancer, the model performed comparable with the doctors' consensus- 72.6%, compared with 70.8%. |
| Shah et al. [25] | 2021 | 4,212 images | HCC, hepatic cysts, liver metastases, hemangioma | The algorithm consisted of a multi-channel deep learning neural network and obtained an accuracy of 98.78% in classifying the lesions, as well as a dice score of 95.7% in locating them. |
| Lee et al. [26] | 2021 | 1,290 images | Hepatic cysts, hemangiomas, liver metastases | The LINA-5 patch achieved the best accuracy of 85.34%, with cyst and hemangioma specificities of 85.47% and 87.73% respectively, and a sensitivity in detecting metastases of 89.93%. |
| Zhou et al. [27] | 2021 | 616 images | HCC, ICC, liver metastases, hepatic cysts, hemangiomas, FNH | Average precision of 82.8%, with 82.5% in detecting whether the tumos is benign or malignant, and 73.4% in detecting the exact type of tumor. |

For abbreviations see Table I

A study performed by Bilello et al. [21] analyzed an algorithm capable of hypodense hepatic lesions automatic diagnosis [21]. Computed tomography scan images from 51 patients were included in the study and the most common hypodense liver lesions, included 22 simple cysts, 11 hemangiomas, 22 metastases, and 1 image containing both a cyst and a hemangioma [21]. The algorithm was programmed to analyze the liver using intensity-based histogram methods for central lesions and liver contour refinement to detect peripheral lesions [21]. Pair-wise lesion classification is applied by support vector machine. The results showed a sensitivity of 80% with 0.8 false positives per section and for 90% sensitivity, the system had 2.2 false positives per section [21].

Yang et al. [22] analyzed the possibility of automatic diagnosis of FLLs using the bag-of-visual-words representations of single and multiple phases method. Contrast-enhanced CT images were used for training and testing the system. The results show that mean average precision can reach more than 90 % and the combined representations of the three enhance phases can improve the mean average precision up to 94.5% [22].

A deep learning model was used by Shi et al. [23] to evaluate a three-phase dynamic contrast enhanced CT protocol for differential diagnosis of HCC from other FLLs [23]. 449 images of FLLs were included in the study and were categorized into HCC and non-HCC groups. Three convolutional dense networks were trained on images scanned with a four-phase CT protocol (precontrast, arterial, portal-venous, and delayed phase). The results showed that a three-phase CT protocol without precontrast showed similar diagnostic accuracy as a four-phase protocol in differentiating HCC from other FLLs [23]. The authors concluded that a multiphase CT protocol for FLLs diagnosis might be optimized by removing the precontrast phase to reduce the radiation dose [23].

Additionally, Gao et al. [24] constructed a deep learning model in order to help differentiate between HCC, ICC, and hepatic metastases [24]. To this extent, the team used images from 723 patients confirmed with these types of cancers, belonging to two different centers, and divided them into a training set and two test sets. The training set consisted of 499 patients from the first center, the first training set of 113 patients from the first center, while the external test set consisted of 111 patients from the second center. Overall, the model achieved an accuracy of 86.2% and AUC of 0.893 on the test set for HCC and ICC, and an accuracy of 72.6%, comparable with the human consensus (70.8%) when differentiating between these two entities. When it comes to the external test set, the model performed with an accuracy of 82.9%. Moreover, it was proven as a great assistance-tool in cases which have been initially misdiagnosed, especially when it comes to the differential diagnosis between ICC and hepatic metastases [24].

Shah et al. [25] used multi-channel deep learning convolutional neural networks in order to diagnose and localize different hepatic lesions on contrast-enhanced CT images, such as HCC, hemangiomas, cysts, and metastases. Overall, 4,212 images were used to train the model, obtaining accuracy of 98.78%, specificity of 98.67%, and specificity of 98.82% respectively, with a dice score for locating the lesions of 95.7% [25]. Lee et al. [26] also used convolutional neural networks to build an algorithm with several patches that would correctly identify and differentiate between various FLLs, such as cysts, hemangiomas, and metastases [26]. The best patch achieved an average accuracy of 85.34%, with the highest sensitivity being for the detection of liver metastases, of 89.93%, while the cyst and hemangioma specificities were 85.47% and 87.73% respectively [26].

Zhou et al. [27] proposed an algorithm based on hierarchical convolutional neural networks, with the aim of performing an automatic diagnosis of FLLs in CT images. To this extent, a total of 616 liver lesions were used and divided into a training and a test set, and the task of the algorithm was to classify the images as either benign or malignant, as well as further suggest the diagnosis (hemangioma, FNH, cyst, HCC, ICC, metastases). The algorithm achieved an overall accuracy of 82.8%, with 82.5% accuracy in detecting whether the tumor was benign or malignant, as well as 73.4% accuracy in diagnosing the exact type of tumor [27].

**Artificial Intelligence based on Magnetic Resonance Imaging**

We found 10 articles analyzing automated diagnosis using MRI images (Table III).

A study performed by Goehler et al. [28] analyzed a system based on a convolutional neural network that was developed to detect liver metastases on MRI and to assess the change in tumor size on consecutive examinations. Kuhn-Munkres algorithm was used for 64 patients with neuroendocrine tumors who performed two consecutive liver MRIs with gadoxetic acid [28]. The results showed that the system was concordant in 91% with the radiologists‘ diagnosis, the sensitivity and specificity was 0.85 (95% confidence interval (95% CI): 0.77; 0.93), respectively 0.92 (95% CI: 0.87; 0.96) [28]. Further, the system was capable of assessing the interval change in tumor burden between two MRI examinations [28].

Images obtained from a multi-phasic MRI were used for the development and validation of a proof-of-concept convolutional neural network that is programmed for automatic diagnosis of FLLs, in a study performed by Hamm et al [29]. The results showed 92% accuracy, a 92% sensitivity,

**Table III.** Studies assessing automated diagnosis of focal liver lesions using magnetic resonance imaging

| Study | Publication Year | Total number of images | Diagnosis | Main Findings |
|---|---|---|---|---|
| Goehler et al. [28] | 2020 | 64 images | Benign and malignant FLLs | Compared with other studies, the automatic system assessed the interval change in the tumor burden between two MRI examinations and had a sensitivity of 0.85 and specificity of 0.92 to classify liver segments as pathological or normal. |
| Hamm et al. [29] | 2019 | 494 images | Simple cyst, cavernous hemangioma, FNH, HCC, ICC, colorectal cancer metastasis. | The study based on a deep learning method showed an accuracy of 92%, sensitivity of 92%, and a 98% specificity, and demonstrated feasibility for automated diagnosis for six common hepatic lesion types. |
| Wang et al. [30] | 2019 | 494 images | Simple cyst, cavernous hemangioma, FNH, HCC, ICC, colorectal cancer metastasis. | The system showed positive predictive value of 76.5%, 82.9% sensitivity and misclassified 12% of lesions. |
| Zhang et al. [31] | 2009 | 320 images | Hepatic cyst, HCC, dysplasia in cirrhosis, cavernous hemangioma, and metastasis. | The artificial neural network method was capable of a training accuracy of 100%, for FLLs classified into five categories. |
| Jansen et al. [32] | 2019 | 271 images | Hepatic cyst, hemangioma, HCC, ICC, liver metastasis. | The images obtained form abdominal magnetic resonance examination, which includes T2-weighted and dynamic contrast enhanced images, were automatically classified with an overall accuracy of 0.77, for all five types of FLLs. |
| Zhen et al. [33] | 2020 | 31,608 images | Hepatic cyst, hemangioma, HCC, liver cirrhosis, cavernous hemangioma, ICC, liver metastasis. | The automatic system based on a deep learning method achieved a performance on par with human experts on classifying FLLs in seven categories. |
| Stollmayer et al. [34] | 2021 | 216 images | FNH, HCC, liver metastases | The deep learning method was used in order to build a 2D and a 3D model as well. The 2D model achieved an AUC of 0.99, 0.96, and 0.995 for the correct identification of the evaluated tumors. The 3D model achieved and AUC of 0.97, 0.9050, and 0.9550 respectively. |
| Alksas et al. [35] | 2021 | 95 patients | Benign and malignant FLLs | The algorithm achieved a sensitivity and specificity of 91.8% and 91.2% respectively. |
| Wang et al. [36] | 2021 | 557 images | Liver cyst, hemangioma, abscess, FNH, HCC, ICC, metastases | The AUC for the two-way classification was 0.969, and for the seven-way one, it reached up to 0.999. Moreover, the overall accuracy was 79.6%. |
| Oestmann et al. [37] | 2021 | 150 lesions | HCC and non-HCC | The average accuracy was 87.3%, with the sensitivities and specificities for the HCC group 92.7% and 82%, while for the non-HCC group were 82% and 92.7% respectively. |

For abbreviations see Table I.

and a 98% specificity [29]. The system displays a sensitivity of 90% for the diagnosis of HCC compared with 70% for radiologists [29]. The authors of the study continued their work with the analysis of a subset of each lesion class and the system misclassified 12% of FLLs [30].

Zhang et al. [31] analyzed the feasibility of a CAD system for FLLs named Liver ANN which classifies liver lesions into five categories using an artificial neural network technique. The automatic system used 320 MRI images obtained from 80 patients. Unfortunately, for each MRI image a radiologist had to delineate a hepatic region of interest, making the system dependent on human intervention. The output of the algorithm is represented by the five hepatic pathologic categories: hepatic cyst, HCC, dysplasia in cirrhosis, cavernous hemangioma, and metastasis. The results showed a training accuracy of 100% and a testing accuracy of 93%. The authors concluded that Liver ANN can provide a second opinion for radiologists [31].

A study performed by Jansen et al. [32] analyzed a system of automatic classification of FLLs using MRI images. Additionally, the system is using also risk factors to obtain a more precise diagnosis [32]. The results showed an overall accuracy for FLLs of 0.77 and a sensitivity/specificity of 0.80/0.78, 0.93/0.93, 0.84/0.82, 0.73/0.56, and 0.62/0.77 for adenoma, cyst, hemangioma, HCC, and metastasis, respectively [32]. The high accuracy obtained by the automatic system can be explained by the fact that the following risk factors were analyzed by the algorithm: hepatic steatosis (only if it was present in the clinical report and the corresponding parameter was assigned with the one value, in absence the zero value). An analog procedure was made for the presence of liver cirrhosis or the presence of a primary tumor [32].

Starting from the fact that dynamic contrast-enhanced MRI provides the most precise diagnosis of hepatic tumors, a study performed by Zhen et al. [33], analyzed the efficiency of a deep learning diagnosis system of FLLs based on convolutional neural networks using enhanced MRI, unenhanced MRI, and relevant clinical information. The results showed that the automatic deep learning-based system is differentiating well malignant from benign FLLs using only unenhanced images (AUC=0.946; 95%CI: 0.914–0.979 vs. AUC=0.951; 95%CI: 0.919–0.982, p=0.664). Further, if the system is combining unenhanced images with clinical data, the performance is considerably improved for classifying malignancies as HCC (AUC=0.985; 95%CI: 0.960–1.000), metastatic tumors (AUC=0.998; 95%CI: 0.989–1.000), and other primary malignancies (AUC=0.963; 95%CI: 0.896–1.000) [33]. The results were compared with the results obtained from pathology examination and the agreement was 91.9%. The authors concluded that sensitivity and specificity of almost every category of FLLs reached the same accuracy compared to three experienced radiologists [33].

Stollmayer et al. [34] used deep learning in order to create two networks, a 2D and a 3D one, respectively, in order to diagnose FNH, HCC, and liver metastases on hepatocyte-specific contrast-enhanced MRI. Thus, 69 patients were included in the study, and a total of 216 images were used in order to train and test the algorithms [34]. Overall, the 2D model performed better than the 3D one, achieving an AUC of 0.99, 0.9664, and 0.96 for the three investigated pathologies (in comparison with 0.97, 0.905, and 0.955 for the 3D model) [34].

Alksas et al. [35] developed a computer-aided diagnosis method in order to correctly identify and classify hepatic tumors on contrast enhanced MRI examinations, based on the LI-RADS classification. Thus, a total of 95 patients were included in the study, and the average sensitivity and specificity for the algorithm were 91.8% and 91.2% respectively. Moreover, the best accuracies were reached in the case of definitely and probably benign tumors (LI-RADS 1 and 2), with 88% and 85%. The lowest accuracy was recorded for the LI-RADS 3 and LI-RADS 5 tumors (78 and 79%) [35].

Additionally, Wang et al. [36] built a model based on convolutional neural networks in order to differentiate various FLLs, by dividing them into benign and malignant, and afterwards, performing a detailed classification depending on the type of the tumors. A total of 557 images were separated into a training and a testing set, and the AUC for the classifications were 0.969 and 0.919, respectively. Moreover, the accuracy for performing the seven-way classification was 79.6% [36]. Another study used convolutional neural networks in order to differentiate between HCC and non-HCC, as well as atypical cases of HCC [37]. The overall accuracy was 87.3%, while the sensitivity and specificity for HCC were 92.7% and 82% respectively, while for the non-HCC group 82% and 92.7% [37].

## DISCUSSIONS

In the field of image analysis, AI has shown promising results in automatic detection, diagnosis, classification, risk stratification, prognosis, and treatment response in connection with a wide variety of organ systems and pathology. Although incipient, applications of AI within hepathology, show significant promise and will likely improve our diagnosis precision of FLLs.

Accurate diagnosis of FLLs is of great importance to improving liver cancer diagnosis, surgery planning and prediction of patient outcome. Nevertheless, human diagnosis is highly expensive, time-consuming, and error-prone to classify the tumors as benign and malignant.

To support developing autonomous AI-based systems for this task, constant efforts have been assigned. Unfortunately, it is still difficult to precisely identify FLLs, which are often diffused, poorly contrasted, and their boundaries are easily confused with healthy liver tissue. Therefore, automatic diagnosis has a great potential to improve the differentiation between malignant and benign liver tumors, which would facilitate the physician's decision-making process and the patients survival rate.

The findings of the current study, show that most available AI-based systems for the automatic diagnosis of FLLs achieve a performance on par with human experienced radiologists, while some models have also achieved better diagnosing and classifying accuracy than the human experts. Using different types of algorithms, the automatic system can distinguish malignant from benign tumors, as well as further classify the type of tumors, and, if the clinical data and laboratory tests are additionally introduced in the system, a higher diagnosis precision is obtained, demonstrating the feasibility of a computed-aided diagnosis of FLLs. Moreover, AI has also proven useful in situations of initial misdiagnosis [24]. Neural networks were the most frequent type of AI-based system used for the automatic diagnosis of FLLs using US images.

In order to better prepare patients for surgical resection, convolutional neural networks have also been used for a more accurate liver segmentation and tumor identification on contrast-enhanced CT scans [38-39]. This method has shown promising results, and further models could provide a great advantage to the clinicians when it comes to managing FLLs. Moreover, there are also studies which have developed certain algorithms based on differentiating FLLs in CEUS, showing promising results as well [40-41].

Our study has several strengths. First, we analyzed all the available imagistic methods for an automatized diagnosis of FLLs using AI-based technologies, including liver ultrasound, CT and MRI. Second, the subject of this systematic review is of major relevance due to the potential of increasing the diagnostic precision of human radiologists independent on the imagistic method used for the diagnosis of FLLs. Third, the COVID-19 curfew, and the serious global human resources shortage makes autonomous diagnosis systems, a fast and affordable solution. Fourth, implementing artificial intelligence algorithms has the potential to improve healthcare, reduce costs, contribute to an evidence-based practice, as well as obtaining a tailored management for the patient [42].

This systematic review has several limitations, which should be acknowledged. The first limitation is the reduced number of subjects included in most studies. International collaboration is recommended to overcome this limitation, and, if a significant number of subjects will be included in future studies, the precision of automatic diagnosis systems will further increase. Secondly, most studies did not specify and classify their sources of error related to the AI-algorithm. Thirdly, studies differed regarding the implemented methods, design, appraisal, and outcomes, and this discrepancy further complicates comparing the results.

The speed of implementation of automatic diagnosis systems in healthcare is directly influenced by challenges such as feasibility, ethical concerns, precision, safety, and overall acceptability. Nevertheless, collaboration between AI-based systems and healthcare professionals still remains a mandatory objective to succeed in such a complex task, and AI cannot replace skilled diagnosticians.

Finally, according to the discussed future challenges, the process of evaluation and comparison of AI techniques used in the autonomous diagnosis of FLLs using medical images, needs further analysis, and adopting a detailed, yet precise criteria for future studies represent an efficient approach to solve this complex issue.

## CONCLUSIONS

We found significant evidence demonstrating that implementing a computer-aided diagnosis for FLLs diagnosis using AI-based applications is currently feasible. This would benefit both the clinician and the patient, as the overall diagnosis accuracy would increase. Moreover, certain AI algorithms have proven beneficial in detecting early-stage hepatic tumors, as well as to better classify initially uncertain types of tumors, leading to a more personalized and precise approach.

**Conflicts of interest:** None to declare.

**Authors' contribution:** S.L.P. gave the idea of this review, searched the literature, analyzed the results, and made a substantial contribution to the writing of the manuscript, D.I.D., S.G. and V.D.B. searched the literature and analyzed the results, G.C. revised the manuscript and made substantial contributions to the methodology of the manuscript, A.M., G.P. analyzed the results, and revised the manuscript. A.F. analyzed the results and made a substantial contribution to the writing of the manuscript. All the authors read and approved the final version of the manuscript.

## REFERENCES

1. Marrero JA, Ahn J, Rajender Reddy K; Americal College of Gastroenterology. ACG clinical guideline: the diagnosis and management of focal liver lesions. Am J Gastroenterol 2014;109:1328-1347. doi:10.1038/ajg.2014.213
2. Boutros C, Katz SC, Espat NJ. Management of an incidental liver mass. Surg Clin North Am 2010;90:699-718. doi:10.1016/j.suc.2010.04.005
3. Bonder A, Afdhal N. Evaluation of liver lesions. Clin Liver Dis 2012;16:271-283. doi:10.1016/j.cld.2012.03.001
4. Gore RM, Pickhardt PJ, Mortele KJ, et al. Management of Incidental Liver Lesions on CT: A White Paper of the ACR Incidental Findings Committee. J Am Coll Radiol 2017;14:1429-1437. doi:10.1016/j.jacr.2017.07.018
5. European Association for the Study of the Liver (EASL). EASL Clinical Practice Guidelines on the management of benign liver tumours. J Hepatol 2016;65:386-398. doi:10.1016/j.jhep.2016.04.001
6. Gatos I, Tsantis S, Spiliopoulos S, et al. A new automated quantification algorithm for the detection and evaluation of focal liver lesions with contrast-enhanced ultrasound. Med Phys 2015;42:3948-3959. doi:10.1118/1.4921753
7. Hwang YN, Lee JH, Kim GY, Jiang YY, Kim SM. Classification of focal liver lesions on ultrasound images by extracting hybrid textural features and using an artificial neural network. Biomed Mater Eng 2015;26 Suppl 1:S1599-S15611. doi:10.3233/BME-151459
8. Marya NB, Powers PD, Fujii-Lau L, et al. Application of artificial intelligence using a novel EUS-based convolutional neural network model to identify and distinguish benign and malignant hepatic masses. Gastrointest Endosc 2021;93:1121-1130.e1. doi:10.1016/j.gie.2020.08.024
9. Mittal D, Kumar V, Saxena SC, Khandelwal N, Kalra N. Neural network based focal liver lesion diagnosis using ultrasound images. Comput Med Imaging Graph 2011;35:315-323. doi:10.1016/j.compmedimag.2011.01.007
10. Virmani J, Kumar V, Kalra N, Khandelwal N. Neural network ensemble based CAD system for focal liver lesions from B-mode ultrasound. J Digit Imaging 2014;27:520-537. doi:10.1007/s10278-014-9685-0
11. Streba CT, Ionescu M, Gheonea DI, et al. Contrast-enhanced ultrasonography parameters in neural network diagnosis of liver tumors. World J Gastroenterol 2012;18:4427-4434. doi:10.3748/wjg.v18.i32.4427

12. Sugimoto K, Shiraishi J, Moriyasu F, Doi K. Computer-aided diagnosis of focal liver lesions by use of physicians' subjective classification of echogenic patterns in baseline and contrast-enhanced ultrasonography. Acad Radiol 2009;16:401-411. doi:10.1016/j.acra.2008.09.018
13. Shiraishi J, Sugimoto K, Moriyasu F, Kamiyama N, Doi K. Computer-aided diagnosis for the classification of focal liver lesions by use of contrast-enhanced ultrasonography. Med Phys 2008;35:1734-1746. doi:10.1118/1.2900109
14. Schmauch B, Herent P, Jehanno P, et al. Diagnosis of focal liver lesions from ultrasound using deep learning. Diagn Interv Imaging 2019;100:227-233. doi:10.1016/j.diii.2019.02.009
15. Acharya UR, Koh JEW, Hagiwara Y, et al. Automated diagnosis of focal liver lesions using bidirectional empirical mode decomposition features. Comput Biol Med 2018;94:11-18. doi:10.1016/j.compbiomed.2017.12.024
16. Xi IL, Wu J, Guan J, et al. Deep learning for differentiation of benign and malignant solid liver lesions on ultrasonography. Abdom Radiol (NY) 2021;46:534-543. doi:10.1007/s00261-020-02564-w
17. Ryu H, Shin SY, Lee JY, Lee KM, Kang HJ, Yi J. Joint segmentation and classification of hepatic lesions in ultrasound images using deep learning. Eur Radiol 2021;31:8733-8742. doi:10.1007/s00330-021-07850-9
18. Tiyarattanachai T, Apiparakoon T, Marukatat S, et al. Development and validation of artificial intelligence to detect and diagnose liver lesions from ultrasound images. PLoS One 2021;16:e0252882. doi:10.1371/journal.pone.0252882
19. Mao B, Ma J, Duan S, Xia Y, Tao Y, Zhang L. Preoperative classification of primary and metastatic liver cancer via machine learning-based ultrasound radiomics. Eur Radiol 2021;31:4576-4586. doi:10.1007/s00330-020-07562-6
20. Massoptier L, Casciaro S. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans. Eur Radiol 2008;18:1658-1665. doi: 10.1007/s00330-008-0924-y
21. Bilello M, Gokturk SB, Desser T, Napel S, Jeffrey RB Jr, Beaulieu CF. Automatic detection and classification of hypodense hepatic lesions on contrast-enhanced venous-phase CT. Med Phys 2004;31:2584-2593. doi:10.1118/1.1782674
22. Yang W, Lu Z, Yu M, Huang M, Feng Q, Chen W. Content-based retrieval of focal liver lesions using bag-of-visual-words representations of single- and multiphase contrast-enhanced CT images. J Digit Imaging 2012;25:708-719. doi:10.1007/s10278-012-9495-1
23. Shi W, Kuang S, Cao S, et al. Deep learning assisted differentiation of hepatocellular carcinoma from focal liver lesions: choice of four-phase and three-phase CT imaging protocol. Abdom Radiol (NY) 2020;45:2688-2697. doi:10.1007/s00261-020-02485-8
24. Gao R, Zhao S, Aishanjiang K, et al. Deep learning for differential diagnosis of malignant hepatic tumors based on multi-phase contrast-enhanced CT and clinical data. J Hematol Oncol 2021;14:154. doi:10.1186/s13045-021-01167-2
25. Shah S, Mishra R, Szczurowska A, Guziński M. Non-invasive multi-channel deep learning convolutional neural networks for localization and classification of common hepatic lesions. Pol J Radiol 2021;86:e440-e448. doi:10.5114/pjr.2021.108257
26. Lee H, Lee H, Hong H, Bae H, Lim JS, Kim J. Classification of focal liver lesions in CT images using convolutional neural networks with lesion information augmented patches and synthetic data augmentation. Med Phys 2021;48:5029-5046. doi:10.1002/mp.15118
27. Zhou J, Wang W, Lei B, et al. Automatic Detection and Classification of Focal Liver Lesions Based on Deep Convolutional Neural Networks: A Preliminary Study. Front Oncol 2021;10:581210. doi:10.3389/fonc.2020.581210
28. Goehler A, Harry Hsu TM, Lacson R, et al. Three-Dimensional Neural Network to Automatically Assess Liver Tumor Burden Change on Consecutive Liver MRIs. J Am Coll Radiol 2020;17:1475-1484. doi:10.1016/j.jacr.2020.06.033
29. Hamm CA, Wang CJ, Savic LJ, et al. Deep learning for liver tumor diagnosis part I: development of a convolutional neural network classifier for multi-phasic MRI. Eur Radiol 2019;29:3338-3347. doi:10.1007/s00330-019-06205-9
30. Wang CJ, Hamm CA, Savic LJ, et al. Deep learning for liver tumor diagnosis part II: convolutional neural network interpretation using radiologic imaging features. Eur Radiol 2019;29:3348-3357. doi:10.1007/s00330-019-06214-8
31. Zhang X, Kanematsu M, Fujita H, et al. Application of an artificial neural network to the computer-aided differentiation of focal liver disease in MR imaging. Radiol Phys Technol 2009;2:175-182. doi:10.1007/s12194-009-0062-5
32. Jansen MJA, Kuijf HJ, Veldhuis WB, Wessels FJ, Viergever MA, Pluim JPW. Automatic classification of focal liver lesions based on MRI and risk factors. PLoS One 2019;14:e0217053. doi:10.1371/journal.pone.0217053
33. Zhen SH, Cheng M, Tao YB, et al. Deep Learning for Accurate Diagnosis of Liver Tumor Based on Magnetic Resonance Imaging and Clinical Data. Front Oncol 2020;10:680. doi:10.3389/fonc.2020.00680
34. Stollmayer R, Budai BK, Tóth A, et al. Diagnosis of focal liver lesions with deep learning-based multi-channel analysis of hepatocyte-specific contrast-enhanced magnetic resonance imaging. World J Gastroenterol 2021;27:5978-5988. doi:10.3748/wjg.v27.i35.5978
35. Alksas A, Shehata M, Saleh GA, et al. A novel computer-aided diagnostic system for accurate detection and grading of liver tumors. Sci Rep 2021;11:13148. doi:10.1038/s41598-021-91634-0
36. Wang SH, Han XJ, Du J, et al. Saliency-based 3D convolutional neural network for categorising common focal liver lesions on multisequence MRI. Insights Imaging. 2021;12:173. doi:10.1186/s13244-021-01117-z
37. Oestmann PM, Wang CJ, Savic LJ, et al. Deep learning-assisted differentiation of pathologically proven atypical and typical hepatocellular carcinoma (HCC) versus non-HCC on contrast-enhanced MRI of the liver. Eur Radiol 2021;31:4981-4990. doi:10.1007/s00330-020-07559-1
38. Alirr OI. Deep learning and level set approach for liver and tumor segmentation from CT scans. J Appl Clin Med Phys 2020;21:200-209. doi:10.1002/acm2.13003
39. Pang S, Field M, Dowling J, Vinod S, Holloway L, Sowmya A. Training radiomics-based CNNs for clinical outcome prediction: Challenges, strategies and findings. Artif Intell Med 2022;123:102230. doi:10.1016/j.artmed.2021.102230
40. Căleanu CD, Sîrbu CL, Simion G. Deep Neural Architectures for Contrast Enhanced Ultrasound (CEUS) Focal Liver Lesions Automated Diagnosis. Sensors (Basel) 2021;21:4126. doi:10.3390/s21124126
41. Hu HT, Wang W, Chen LD, et al. Artificial intelligence assists identifying malignant versus benign liver lesions using contrast-enhanced ultrasound. J Gastroenterol Hepatol 2021;36:2875-2883. doi:10.1111/jgh.15522
42. Aboueid S, Liu RH, Sabesan P. The cost-effectiveness of algorithms and artificial intelligence applied in health care: A scoping review research protocol. JMIR Preprints 24/11/2019:17172. doi:10.2196/preprints.17172