Original paper

Is artificial intelligence able to discriminate emergencies?

Raquel Bernal Calmarza[a], Ana Valer Martínez[b], María Celada Suárez[c], Sara Calmarza Delgado[d], Elena Calmarza Delgado[d]

[a]Paediatrician. CS Quince de mayo. Madrid. Spain • [b]Family physician. CS Tarazona. Tarazona. Zaragoza. Spain • [c]Medical resident in family medicine. Hospital Universitario Miguel Servet. Zaragoza. Spain • [d]Nurse. Hospital Ernest Lluch Martin. Calatayud. Zaragoza. Spain.

Date of online publication: 31-october-2024

Raquel Bernal Calmarza: raquel3433@gmail.com

Abstract

Introduction: in paediatrics, high-frequency emergency department use is defined as repeated emergency visits for reasons that do not require urgent attention or could be managed at a different level of care. Several factors may be associated with this phenomenon, such as socioeconomic, cultural or psychological factors. Its impact on the health care system is significant. Artificial intelligence (AI) has the potential of reducing high-frequency use.

Methodology: we assessed the agreement between the information for 101 diseases common in children provided by Gemini AI, a free and open-access service, and the current scientific evidence. We used the adjusted kappa coefficient in this analysis.

Results: the AI provided responses for all of the 101 diseases considered in the analysis. The kappa coefficient was 0.857 (95% CI, 0.002) for the identification of the disease, 0.888 (95% CI, 0.003) for the identification of warning signs, 0.876 (95% CI, 0.005) for establishing the need to visit the emergency department and 0.915 (95% CI, 0.003) for the appropriate recommendation of measures to be taken.

Conclusions: the text-based artificial intelligence exhibited substantial agreement with protocols used for identification of diseases based on symptoms, and near-perfect agreement for determining the need to visit the emergency department, identifying warning signs and providing therapeutic recommendations. The level of agreement was higher for common diseases and children aged more than 3 months.

Key words:
- Artificial intelligence
- Emergencies
- Diagnosis

¿Puede la inteligencia artificial discriminar urgencias?

Resumen

Introducción: la hiperfrecuentación en Pediatría se define como la asistencia repetida a urgencias por motivos que no requieren atención urgente o podrían ser tratados en otro nivel asistencial. Los factores que contribuyen son diversos, y pueden incluir factores socioeconómicos, culturales y psicológicos. El impacto en el sistema de salud es significativo. La inteligencia artificial (IA) tiene el potencial de ser una herramienta eficaz para reducir la hiperfrecuentación.

Metodología: se analiza la concordancia entre la información aportada por la inteligencia artificial Gemini, de acceso libre y gratuito, para 101 enfermedades frecuentes en la infancia, en comparación con la evidencia disponible. Se analiza con el coeficiente kappa ajustado.

Resultados: de las 101 patologías analizadas, la IA dio una respuesta en todas ellas. Se obtuvo un reconocimiento de la patología con un coeficiente kappa de 0,857 +/- 0,002, un reconocimiento de los signos de alarma de 0,888 +/- 0,003, una adecuación de la necesidad de acudir a urgencias de 0,876 +/- 0,005 y una adecuación de las medidas a tomar de 0,915 +/- 0,003.

Conclusiones: la inteligencia artificial basada en texto tiene una concordancia muy buena respecto a los protocolos para reconocer patologías a partir de síntomas, y muy buena para valorar la necesidad de visita a un servicio de urgencias, la valoración de los signos de alarma y las recomendaciones terapéuticas. Esta concordancia es mayor en niños mayores de tres meses de edad y para patologías comunes.

Palabras clave:
- Diagnóstico
- Inteligencia artificial
- Urgencias

How to cite this article: Bernal Calmarza R, Valer Martínez A, Celada Suárez M, Calmarza Delgado S, Calmarza Delgado E. ¿Puede la inteligencia artificial discriminar urgencias? Rev Pediatr Aten Primaria. 2024;26:351-60. https://doi.org/10.60147/dce30dee

INTRODUCTION

High-frequency use of health care services is a growing problem in paediatrics that can have a deleterious impact on care quality and available health care resources. In the emergency care setting, high-frequency use is defined as making recurrent visits to the emergency department for reasons that do not require urgent care or could be managed in a different setting or level of care.

The factors that contribute to high-frequency use of paediatric emergency care services are varied and may include socioeconomic, cultural and psychological factors.[1] Possible socioeconomic factors include lack of access to primary care, poverty or immigrant status. Among cultural factors, specific beliefs and practices may contribute significantly to the excessive use of emergency care. The psychological factors include anxiety, stress and lack of trust in primary care services.

High-frequency users have a significant impact on the health care system. They are a small proportion of the paediatric population, but give rise to a large volume of visits.[2] This can overburden primary and emergency care services and increase health care costs.

In paediatric care, high-frequency use is a complex problem that needs to be addressed from a multidisciplinary approach. Possible strategies include education of patients and parents, improving access to primary care and the development of specific programmes for the management of high-frequency users.

Artificial intelligence (AI) is revolutionising medicine, with applications in multiple areas, ranging from diagnosis and treatment of diseases to research and development of new drugs.[3] Several studies have shown that AI has the potential to be an effective tool to reduce high-frequency use.[4] AI can be used to assist patient triage according to the level of urgency and to provide information and support to patients and their parents.[5]

Although there is evidence that supports the usefulness of AI in assisting triage in paediatric emergency departments, there is no evidence on its performance in supporting parental decision-making regarding the need to visit the emergency department.[6]

The aim of our study was to measure the level of agreement between the results obtained through Gemini artificial intelligence (formerly known as Bard) and the best available evidence for 101 common and significant diseases in terms of four variables: recognition of the disease, identification of warning signs, need to visit emergency department and measures to be taken at home.

One of the secondary objectives was to assess the agreement between the results yielded by the Gemini AI and the best available evidence for 101 common and significant diseases in terms of the same four variables (recognition of the disease, correction of warning signs, need to visit emergency department and measures to be taken at home) for each disease category. We also aimed to identify factors that could affect the precision and accuracy of diagnosis in the Gemini AI.

MATERIAL AND METHODS

We conducted an observational study comparing the results obtained through prompts issued to the Gemini AI model with the best available scientific evidence, defined as the protocols of the Sociedad Española de Urgencias Pediátricas (SEUP, Spanish Society of Paediatric Emergency Medicine)[7] and the algorithms of the Asociación Española de Pediatría de Atención Primaria (AEPap, Spanish Association of Primary Care Paediatrics).[8]

We used the Gemini artificial intelligence, a machine-learning model trained on a massive dataset of text and code, which has the ability to generate text in response to prompts to providing informational answers in Spanish.

Gemini is one of the world's largest language models, with 137 billion parameters. This allows it to learn complex language patterns and relationships. Furthermore, it has access to a massive dataset of text and code, allowing it to learn about a broad range of topics. It can generate high-qual-

ity, grammatically correct and coherent text and provide informative answers to questions, even if they are open-ended. It is an open-access language model that is free for users, so it is easy for anyone with Internet access to use it. This tool has substantial potential for reaching a significant portion of the population.

In the analysis of the results, if the answer provided by the AI coincided completely with the corresponding content in the protocols, we coded it as 'right', and if the answer did not coincide with the protocols or was incomplete, we coded it as 'wrong'. The level of agreement was calculated as the proportion of right answers in relation to the total number of questions submitted to the AI.

If the response of the AI was "es necesario que acuda a urgencias inmediatamente" (you need to go to the emergency department immediately) or "necesita una valoración médica inmediata" (you need immediate medical assessment), it was interpreted as "need to visit the emergency department", and if the response was "debe acudir al médico" (you should go to the doctor) or "lleve a su hijo al médico" (take you child to the doctor) it was interpreted as "no need to visit the emergency department".

We asked the AI about 101 diseases, selected based on their severity and frequency in the months ranging from December 2023 to January 2024. The study was based on previously healthy patients who did not require chronic medication.

Table 1 presents the symptoms of each disease about which we submitted questions to the AI. For each symptom (unless it applied only to a specific age group, in which case it is so noted), we asked about 2 ages: 3 months for infants, and 4 years for young children. In the case of diseases defined in a specific age range (limp in children aged 6-8 years, limp in adolescents) we asked about the specified age.

For 5 diseases, we asked the AI in regard to all age groups (by month between ages 1 and 24 months and by year for ages 2 to 24 years), and found no differences in the responses of the AI in relation to age. For this reason, we decided to limit questions to age 3 months in reference to infants and age 4 years in reference to young children (early childhood and school age).

We used the Cohen kappa correlation coefficient to measure the level of agreement, which was adjusted for chance in the case of dichotomous variables (need/no need to visit emergency department). For variables that were not dichotomous (recognition of disease, appropriate identification of warning signs and measures to take at home), due to the lack of previous data, we calculated the correlation coefficient without adjusting for chance, using an expected agreement by chance at 0.10 for the calculation of confidence intervals, although the actual probability is likely lower. For diseases for which there are no measures to be taken at home because they always require immediate medical attention, we did not analyse the variables "concerning the assessment of warning signs and the measures to be taken at home. The correlation coefficients and confidence intervals were calculated with the software package SPSS version 28.

We considered the agreement near-perfect if the coefficient was greater than 0.8, substantial if it was between 0.6 and 0.8, moderate if it was between 0.4 and 0.5, fair if it was between 0.2 and 0.4 and poor if it was less than 0.2. If the confidence interval extended across agreement levels, we included both levels in the results.

RESULTS

We analysed responses for 101 diseases, selected based on their frequency and severity. For each disease, we analysed the following variables: recognition of the disease based on the submitted symptoms (Table 1), appropriateness of the warning signs identified by the AI, assessment (correct or incorrect) of the need to visit the emergency department and measures to be taken at home.

The AI assessed the need to seek medical care for each of the symptoms.

Table 2 presents the overall results, Table 3 the results for the warning signs by age group and Table 4 the results by disease category.

Table 1. Symptoms contributed to artificial intelligence	
Condition	**Symptom(s)**
Abdominal pain	Tummy ache
Acute bronchiolitis	Breathing difficulty
Acute dysphagia	Unable to swallow for several hours
Acute myositis	Calf pain and inability to walk
Acute scrotum	Testicular pain
Acute sinusitis	Nasal discharge and headache with or without fever
AGE	4 vomiting and 4 diarrhoea
Allergy	Hives after ingestion
Allergy	Vomiting after ingestion
Allergy	Difficulty breathing after ingestion
Anxiety	Nervousness
Arthritis, monoarticular	Knee pain and swelling
Arthritis, polyarticular	Pain and swelling in hands, wrists and knees
Asthma	Respiratory distress
Asthma	Cough
Asthma	Chest tightness
Bite	Dog/cat bite
Bradycardia	Slow heart rate
Breath-holding spell	Cyanosis and fainting with crying in an infant
Bullous rash	Skin blisters
Burn	Boiling water burn
Burn	Frying pan burns
Cardiorespiratory arrest	Loss of consciousness with absence of breathing
Cellulite	Erythematous swelling around a wound
Chest pain	Chest pain
CMPA	Vomiting upon introducing cow's milk
CMPA	Blood in stools upon introducing cow's milk
CMPA	Irritability upon introducing cow's milk
Cold	Cough, nasal discharge and fever
Coma/decreased consciousness	Loss of consciousness with impaired breathing
Conjunctivitis	Red eye
Constipation in infants	4 days without bowel movements
Constipation in preschool-/school-aged children	4 days without bowel movements
Cutaneous mycosis	Red patch on one foot
Cyanosis	Blue or purple discoloration of lips
Dental pain	Pain in a tooth
Dental phlegmon	Toothache with swelling of the face
Diarrhoea	6 bowel movements in the past day
Diplopia	Seeing double
Epistaxis	Nosebleeds
Febrile respiratory infection	Cough, mucus and fever up to 39 °C
Fever	Fever
Foreign body, ear	Sticking a coin/chickpea in the ear
Foreign body, gastrointestinal	Swallowing a coin
Foreign body, gastrointestinal	Swallowing a battery
Foreign body, nose	Sticking a coin/chickpea up the nose
Foreign body, ocular	Getting sand in the eye

AGE: acute gastroenteritis; **AOM:** acute otitis media; **CMPA:** cow's milk protein allergy; **GOR:** gastro-oesophageal reflux; **SIDS**: sudden infant death syndrome; **TBI:** traumatic brain injury; **UTI**: urinary tract infection.

Continued of next page

Table 1. Symptoms contributed to artificial intelligence (*continuation of the previous page*)	
Condition	**Symptom(s)**
Foreign body, respiratory	Swallowing a coin and difficulty breathing
Foreign body, throat	Stuck fish bone
GOR in infant	Pours milk at all feedings
GOR in older child	Chest pain
GOR in older child	Sensation of food coming back up
Gross haematuria	Blood in urine
Haematemesis	Vomiting blood
Headache	Headache
Impetigo	Yellow crusts on the skin
Inborn error of metabolism	Vomiting in infants
Inborn error of metabolism	Seizures in infants
Inborn error of metabolism	Decreased level of consciousness or muscle tone in infants
Increased intracranial pressure	Severe headache
Increased intracranial pressure	Seeing double/paralysis of one side of the face
Insect bite	Insect bite
Irritability	Inconsolable crying
Laryngitis	Hoarse cough
Laryngitis	Stridor
Leukaemia	Fatigue
Leukaemia	Frequent bruising
Leukaemia	Difficult to control bleeding
Leukocoria	Reflection of flash light is not red
Limp in adolescents	Limp
Limp in children aged 6-8 years	Limp
Limp in toddler/preschooler	Limp
Lower gastrointestinal bleeding	Bloody stools
Lymphadenopathy	Lump in the neck or groin
Mumps	Swelling of the face
Neonatal jaundice	Yellowish skin colour in infants
Nephrotic syndrome	Eyelid swelling
Non-neonatal jaundice	Yellow eyes in children
Otalgia	Earache
Palpitations	Stabbing chest pain
Paronychia	Painful red swelling around nail
Pharyngitis/tonsillitis	*Fever and sore throat*
Physical abuse	Bruising on thighs
Physical abuse	Decreased level of consciousness in infant (unresponsive)
Physical abuse	Cigarette burns
Poisoning	Accidental toxic substance ingestion
Rash, febrile maculopapular	Red rashes with fever
Rash, purpuric	Red spots on the skin
Rash, vesicular	Vesicles on the skin
Respiratory allergy	Rhinitis
Respiratory allergy	Conjunctivitis
Scabies	Itching
Scabies	Itchy and blotchy skin
Seizure, afebrile	Seizure without fever
Seizure, febrile	Seizure and fever

AGE: acute gastroenteritis; **AOM:** acute otitis media; **CMPA:** cow's milk protein allergy; **GOR:** gastro-oesophageal reflux; **SIDS**: sudden infant death syndrome; **TBI:** traumatic brain injury; **UTI**: urinary tract infection.

Continued of next page

Table 1. Symptoms contributed to artificial intelligence (*continuation of the previous page*)	
Condition	**Symptom(s)**
Seizure, partial	Myoclonus, unilateral arm jerking
Sexual abuse	Lesions in vulvar region
Sexual abuse	Lesions in anal region
Shock	Tachycardia (very fast heartbeat) and pallor
SIDS	Loss of consciousness with absence of breathing
Stye	Eyelid lump
Suicidal ideation	Child says he wants to commit suicide
Suppurative AOM	Ear discharge
Syncope	Loss of consciousness
Tachycardia	Fast heart rate
TBI in infant	Fall from changing table
Tics	Involuntary movements of the hands or face
Torticollis	Neck pain and difficulty moving the neck
Torticollis with fever	Pain/difficulty moving neck with fever
Trauma, abdominal	Abdominal pain after falling off a bicycle
Trauma, ankle	Ankle pain after a fall
Trauma, dental	Toothache after a fall
Trauma, forearm	Wrist pain after fall
Trauma, high-energy	Fall from height/car accident
Type 1 diabetes	Frequent urination
Type 1 diabetes	High intake with weight loss
Urticaria	Skin rashes
UTI	Burning sensation in passing urine
UTI, febrile in older child	Burning sensation in passing urine
Vertigo	Dizziness and spinning of objects
Vomiting	Vomiting
Wounds	Injury from a fall

AGE: acute gastroenteritis; **AOM:** acute otitis media; **CMPA:** cow's milk protein allergy; **GOR:** gastro-oesophageal reflux; **SIDS**: sudden infant death syndrome; **TBI:** traumatic brain injury; **UTI**: urinary tract infection.

Table 2. Overall results		
Variable	**Kappa coefficient**	**Correlación**
Recognition of the disease	0.857 (0.002)	Near perfect
Correct warning signs	0.888 (0.003)	Near perfect
Need to go to emergency room	0.876 (0.005)	Near perfect
Measures to be taken	0.915 (0.003)	Near perfect

Table 3. Age-adjusted analysis, warning signs		
Age	**Kappa coefficient**	**Correlación**
< 3 months (infant)	0.667 (0.053)	Substantial
≥ 3 months (child)	0.938 (0.004)	Near perfect

Table 4. Results by disease group

Respiratory diseases: asthma, bronchiolitis, upper respiratory tract infection with and without fever, cyanosis		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	1 (0.045)	Near perfect
Correct warning signs	1 (0.053)	Near perfect
Need to go to emergency room	0.733 (0.102)	Substantial
Measures to be taken	0.778 (0.064)	Substantial

Gastrointestinal diseases: neonatal jaundice, non-neonatal jaundice, acute dysphagia, gastro-oesophageal reflux in infant, gastro-oesophageal reflux in older child, haematemesis, lower gastrointestinal bleeding, diarrhoea, vomiting, constipation in infant and constipation in older child		
Variable	kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.933 (0.021)	Near perfect
Correct warning signs	0.909 (0.024)	Near perfect
Need to go to emergency room	1 (0.04	Near perfect
Measures to be taken	0.909 (0.026)	Near perfect

Dermatological diseases and rashes: maculopapular rash with fever, vesicular rash, bullous rash, purpuric rash, impetigo, cellulitis, paronychia, scabies and cutaneous fungal disease		
Variable	kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.636 (0.029)	Substantial/Near perfect
Correct warning signs	0.8 (0.032)	Substantial
Need to go to emergency room	0.733 (0.051)	Substantial
Measures to be taken	0.8 (0.032)	Substantial/Near perfect

Surgery and trauma: high-energy trauma, abdominal trauma, ankle and foot trauma, forearm trauma, burn, bite, insect bite, wound, gastrointestinal foreign body, respiratory foreign body, torticollis, dental pain		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	1 (0.019)	Near perfect
Correct warning signs	1 (0.026)	Near perfect
Need to go to emergency room	1 (0.031)	Near perfect
Measures to be taken	1 (0.019)	Near perfect

Neurologic diseases: irritability, coma, syncope, diplopia, SIDS, increased intracranial pressure, headache, tics, febrile seizure, afebrile seizure, partial seizure, breath-holding spell, acute myositis		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.928 (0.023)	Near perfect
Correct warning signs	1 (0.035)	Near perfect
Need to go to emergency room	1 (0.042)	Near perfect
Measures to be taken	1 (0.032)	Near perfect

Oncological diseases: leukocoria, leukaemia		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.25 (0.032)	Weak
Need to go to the emergency room	1 (0.267)	Substantial /Near perfect

Allergies: CMPA; food allergy; respiratory allergy		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.25 (0.040)	Weak
Correct warning signs	1 (0.162)	Near perfect
Need to go to emergency room	1 (0.174)	Near perfect
Measures to be taken	0.667 (0.107)	Weak/Substantial

AGE: acute gastroenteritis; **AOM:** acute otitis media; **CMPA:** cow's milk protein allergy; **ENT:** ear, nose, throat; **GOR:** gastro-oesophageal reflux; **SIDS:** sudden infant death syndrome; **TBI:** traumatic brain injury; **UTI**: urinary tract infection

Continued of next page

Table 4. Results by disease group (*continuation of the previous page*)

Cardiovascular diseases: chest pain, palpitations, bradycardia, tachycardia		
Variable	kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	1 (0.080)	Near perfect
Need to go to emergency room	0.333 (0.129)	Weak/Moderate
Measures to be taken	1 (0.107)	Near perfect

Infectious diseases: lymphadenopathy, torticollis with fever, dental phlegmon, acute mumps, fever without source, limp in preschool-aged children, limp in school-aged children, limp in adolescents		
Variable	kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.75 (0.040)	Substantial
Correct warning signs	0.583 (0.026)	Moderate
Need to go to emergency room	0.428 (0.073)	Weak/Moderate
Measures to be taken	0.833 (0.053)	Substantial/Near perfect

Ophthalmological and ENT diseases: ocular foreign body, conjunctivitis, stye, foreign body in nose, foreign body in ear, otalgia, suppurative otitis, acute pharyngitis, acute laryngitis, acute sinusitis, foreign body in pharynx, vertigo		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	1 (0.023)	Near perfect
Correct warning signs	1 (0.024)	Near perfect
Need to go to emergency room	0.795 (0.039)	Substantial/Near perfect
Measures to be taken	0.923 (0.023)	Near perfect

Metabolic and endocrine diseases: inborn errors of metabolism and type 1 diabetes		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.8 (0.064)	Substantial/Near perfect
Need to go to emergency room	1 (0.080)	Near perfect

Genitourinary diseases: febrile UTI, febrile UTI in older children, nephrotic syndrome, macroscopic haematuria		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.75 (0.080)	Substantial/Near perfect
Correct warning signs	1 (0.080)	Near perfect
Need to go to emergency room	1 (0.129)	Near perfect
Measures to be taken	1 (0.080)	Near perfect

Rheumatologic diseases: monoarticular and polyarticular arthritis		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.5 (0.162)	Weak/Substantial
Correct warning signs	1 (0.162)	Near perfect
Need to go to emergency room	1 (0.267)	Substantial/Near perfect
Measures to be taken	1 (0.162)	Near perfect

Psychiatric disorders, maltreatment and poisoning: physical abuse, sexual abuse, accidental poisoning, anxiety and suicidal ideation		
Variable	Kappa coefficient (95% confidence interval)	Agreement
Recognition of the disease	0.7 (0.035)	Substantial
Correct warning signs	1 (0.045)	Near perfect
Need to go to emergency room	0.8788 (0.030)	Near perfect
Measures to be taken	1 (0.040)	Near perfect

AGE: acute gastroenteritis; **AOM:** acute otitis media; **CMPA:** cow's milk protein allergy; **ENT:** ear, nose, throat; **GOR:** gastro-oesophageal reflux; **SIDS:** sudden infant death syndrome; **TBI:** traumatic brain injury; **UTI**: urinary tract infection

DISCUSSION

Although the level of agreement of the AI for the recognition of diseases was 0.86, it proved more effective in identifying warning signs, with a kappa coefficient of 0.89, and in determining whether or not there was a need to visit the emergency department, with a kappa coefficient of 0.88 overall, in addition to giving advice to parents, with a kappa coefficient of 0.91.

Although the overall agreement was substantial, there were large differences in the stratified analysis. While the level of agreement was near-perfect in establishing the diagnosis and the level of urgency common diseases (respiratory, gastrointestinal, ophthalmological and otorhinolaryngological diseases and trauma), the obtained kappa coefficients were lower, although still good, for less prevalent conditions or those with less specific symptoms (endocrine, metabolic, cardiovascular, oncological and rheumatological diseases, psychiatric disorders and suspected abuse).

We ought to specifically comment on dermatological complaints which, on account of their very nature, are difficult to describe with words. In spite of this, the AI achieved a substantial level of agreement in the assessment of warning signs and of the need to visit the emergency department. Something to be considered is the possibility of submitting photographs to AI to increase its diagnostic yield.

When it came to warning signs, AI exhibited near-perfect agreement in children aged more than 3 months. In infants under 3 months, the agreement in clinical warning signs decreased considerably (from 0.94 to 0.67) due to the nonspecificity of symptoms in infants and the fact that, due to their young age, these patients are at increased risk of complications. Thus, when it comes to infants under 3 months, AI cannot be considered a reliable enough instrument to recommend its widespread application.

We did not find any other articles in the literature analysing the effectiveness of artificial intelligence in detecting warning signs compared to the use of protocols and scientific evidence, we were unable to compare our findings with those of previous studies.

We may conclude that AI is a useful tool for classifying symptoms as urgent versus less urgent in children older than 3 months, but it should not replace medical consultation, as doing so could result in missing diseases that, although not requiring immediate medical care, may be serious and difficult to detect, such as oncological diseases or child abuse.

CONCLUSION

The text-based AI performed with a substantial level of agreement with respect to paediatric manuals and protocols commonly used to identify diseases based on symptoms, and near-perfect agreement with existing paediatric protocols in determining the need to visit an emergency department, assessing warning signs and making therapeutic recommendations. We found the highest levels of overall agreement for respiratory, gastrointestinal, trauma/surgical, genitourinary, neurologic, ophthalmological and otorhinolaryngological conditions. The correlation coefficients for these groups were greater than 0.7 in every analysed category.

On the other hand, we found the lowest levels of agreement in the recognition of oncological, allergic and rheumatological diseases, although the AI was effective in recognising the warning signs requiring a visit to the emergency department for all these conditions. In addition, the agreement of AI with standard guidelines and protocols decreased significantly for patients aged less than 3 months. Further studies are needed to assess the performance of artificial intelligence compared to the judgment of a health care professional.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare in relation to the preparation and publication of this article.

AUTHORSHIP

All authors contributed equally to the development of the published manuscript.

ABBREVIATIONS

AEPap: Asociación Española de Pediatría de Atención Primaria • **AI:** artificial intelligence • **SEUP:** Sociedad Española de Urgencias Pediátricas.

REFERENCES

1. Martín R, Sánchez M, Teruel C. Aspectos psicosociales relacionados con la hiperfrecuentación en las consultas pediátricas de atención primaria. An Pediatr (Barc). 2019;90(1):26-31. https://doi.org/10.1016/j.anpedi.2018.03.008
2. Mata D, Fernández M, Cobo MV. Hiperfrecuentación en Pediatría: ¿Se trata de una verdadera necesidad? Bol Pediatr. 2015;55(234):243-6.
3. Mancera J, Muñoz F, Paniagua F, Gómez F, Fernández ML, Blanca FJ. Problemas de salud y factores determinantes del número de visitas a demanda en pacientes hiperutilizadores de un centro de salud. Aten Primaria. 2001;27(9):658-62. https://doi.org/10.1016/S0212-6567(01)78876-2
4. Liang H, Tsui BY, Ni H, Valentim CCS, Baxter SL, Liu G, *et al*. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence. Nat Med. 2019;25(3):433-8. https://doi.org/10.1038/s41591-018-0335-9
5. Ramgopal S, Sanchez-Pinto LN, Horvat CM, Carroll MS, Luo Y, Florin TA. Artificial intelligence-based clinical decision support in pediatrics. Pediatr Res. 2023;93(2):334-41. https://doi.org/10.1038/s41390-022-02226-1
6. Jiménez Alés R. Inteligencia artificial. Desafíos y preocupaciones. Rev Pediatr Aten Primaria. 2023;25:205-10.
7. Sociedad Española de Urgencias de Pediatría (SEUP). Protocolos diagnósticos y terapéuticos en urgencias de pediatría. (4th ed.); 2019 [online] [accessed 29/10/2024]. Available at https://seup.org/wp-content/uploads/2024/04/0_Primeras_4ed.pdf
8. Asociación Española de Pediatría de Atención Primaria. Algoritmos AEPap [online] [accessed 29/10/2024]. Available at https://algoritmos.aepap.org/