Appendix:

eTable 1: Scoping Reviews (PRISMA-ScR) Checklist

eTable 2: Database Search Terms and Queries

eTable 3: Language Foundation Models Used in the Included Studies

eTable 4: Multimodal Foundation Models Relevant to Emergency and Critical Care

eTable 5: Vision and Audio Foundation Models Relevant to Emergency and Critical Care

Appendix eTable 1: Scoping Reviews (PRISMA-ScR) Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	1
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	3
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	3-4
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	N/A
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	4
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	4
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	4

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	4
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	4-5
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	4-5
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	N/A
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	N/A
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	6
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	7-9
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	N/A
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	10-11
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	11-12
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence	12-13

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
		available), link to the review questions and objectives, and consider the relevance to key groups.	
Limitations	20	Discuss the limitations of the scoping review process.	14-155
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	16
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	N/A

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).

‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.

§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. Ann Intern Med. 2018;169:467–473. doi: 10.7326/M18-0850.

Appendix eTable 2: Database Search Terms and Queries

Database	Queries	Number of Retrieved Studies
PubMed	("Acute care" OR "Emergency care" OR "Critical care" OR "prehospital" OR "pre-hospital" OR "Emergency" OR "Intensive care" OR "Intensive Care Unit" OR "ICU" OR "SICU" OR "MICU" OR "A&E" OR "ED") AND ("artificial intelligence" OR "machine learning" OR "AI" OR "ML" OR "neural network" OR "neural networks" OR "data mining" OR "pattern recognition" OR "computer reasoning" OR "machine intelligence" OR "deep learning") AND ("foundation model" OR "foundational model" OR "foundation models" OR "pre-train" OR "pretraining" OR "pretrain" OR "pre-trained model" OR "pretrained model" OR "large language model" OR "large language models" OR "LLM" OR "LLMs" OR "general-purpose model" OR "generalizable AI")	339
Web of Science	TS = (("Acute care" OR "Emergency care" OR "Critical care" OR "prehospital" OR "pre-hospital" OR "Emergency" OR "Intensive care" OR "Intensive Care Unit" OR "ICU" OR "SICU" OR "MICU" OR "A&E" OR "ED") AND ("artificial intelligence" OR "machine learning" OR "AI" OR "ML" OR "neural network" OR "neural networks" OR "data mining" OR "pattern recognition" OR "computer reasoning" OR "machine intelligence" OR "deep learning") AND ("foundation model" OR "foundational model" OR "foundation models" OR "pre-train" OR "pretraining" OR "pretrain" OR "pre-trained model" OR "pretrained model" OR "large language model" OR "large language models" OR "LLM" OR "LLMs" OR "general-purpose model" OR "generalizable AI"))	198
Embase	('acute care' OR 'emergency care' OR 'critical care' OR 'prehospital' OR 'pre-hospital' OR 'emergency' OR 'intensive care' OR 'intensive care unit' OR 'icu' OR 'sicu' OR 'micu' OR 'a&e' OR 'ed') AND ('artificial intelligence' OR 'machine learning' OR 'ai' OR 'ml' OR 'neural network' OR 'neural networks' OR 'data mining' OR 'pattern recognition' OR 'computer reasoning' OR 'machine intelligence' OR 'deep learning') AND ('foundation model' OR 'foundational model' OR 'foundation models' OR 'pre-train' OR 'pretraining' OR	480

	'pretrain' OR 'pre-trained model' OR 'pretrained model' OR 'large language model' OR 'large language models' OR 'llm' OR 'llms' OR 'general-purpose model' OR 'generalizable ai')	
Scopus	TITLE-ABS-KEY (("Acute care" OR "Emergency care" OR "Critical care" OR "prehospital" OR "pre-hospital" OR "Emergency" OR "Intensive care" OR "Intensive Care Unit" OR "ICU" OR "SICU" OR "MICU" OR "A&E" OR "ED") AND ("artificial intelligence" OR "machine learning" OR "AI" OR "ML" OR "neural network" OR "neural networks" OR "data mining" OR "pattern recognition" OR "computer reasoning" OR "machine intelligence" OR "deep learning") AND ("foundation model" OR "foundational model" OR "foundation models" OR "pre-train" OR "pretraining" OR "pretrain" OR "pre-trained model" OR "pretrained model" OR "large language model" OR "large language models" OR "LLM" OR "LLMs" OR "general-purpose model" OR "generalizable AI"))	402
CINAHL	(TI OR AB) ("Acute care" OR "Emergency care" OR "Critical care" OR "prehospital" OR "pre-hospital" OR "Emergency" OR "Intensive care" OR "Intensive Care Unit" OR "ICU" OR "SICU" OR "MICU" OR "A&E" OR "ED") AND (TI OR AB) ("artificial intelligence" OR "machine learning" OR "AI" OR "ML" OR "neural network" OR "neural networks" OR "data mining" OR "pattern recognition" OR "computer reasoning" OR "machine intelligence" OR "deep learning") AND (TI OR AB) ("foundation model" OR "foundational model" OR "foundation models" OR "pre-train" OR "pretraining" OR "pretrain" OR "pre-trained model" OR "pretrained model" OR "large language model" OR "large language models" OR "LLM" OR "LLMs" OR "general-purpose model" OR "generalizable AI")	24

Overview of Foundation Model Types Included in This Review

In the included studies, FMs could be broadly categorized into four main types: language foundation models, multimodal foundation models, vision foundation models, and audio foundation models. Among these, nearly all models applied directly in ECC tasks were language models. Other FM families, such as multimodal models (e.g., BLIP, CLIP), vision models (e.g., ViT), and audio models (e.g., Wav2Vec 2.0) did not appear in ECC applications within our included studies but are presented here for conceptual completeness. This classification reflects both the structural design and functional direction of these models, as well as their compatibility with different clinical data modalities encountered in ECC settings.

Appendix eTable 3: Language Foundation Models Used in the Included Studies

Overview of language foundation models applied in emergency and critical care. This table summarizes the language FMs used in the included studies, including their structural definitions, representative model series, specific variants applied in ECC tasks, and associated clinical applications. Language models constituted the vast majority of FM deployments in the reviewed ECC literature.

Category	**Category Definition**	**Representative Model Series**	**Model Series Definition**	**Model Variants applied in ECC**	**Representative Applications in ECC**
Language Foundation Model	Language models are FMs that use transformer neural network architectures and are	Generative Pretrained Transformer (GPT)	GPT is an autoregressive language model trained via generative pretraining on large-scale unlabeled text	BioGPT, GPT-3.5, GPT-3.5-turbo, GPT-4, GPT-4o, GPT-4o Mini	Applied in 71.43% (35/49) of studies. ● Pediatric ED diagnosis evaluation [58] ● ICU discharge summary generation [88] ● Neonatal ICU intervention

	pretrained using self-supervised learning on extensive text corpora to capture complex patterns in language and perform a broad range of language understanding or generation tasks without requiring task-specific tuning [46,147–150].		corpora and capable of performing a wide range of downstream tasks through in-context learning, without requiring task-specific fine-tuning[46,151].		recommendation assessment [77] ● Emergency Severity Index (ESI)-based triage acuity prediction [94] ● ICU data extraction and profiling [16]
		Bidirectional Encoder Representations from Transformers (BERT)	BERT is a deep bidirectional language representation model based on the Transformer encoder architecture, pre-trained using masked language modeling and next sentence prediction to capture contextual dependencies from unlabeled text [47].	RoBERTa, BioBERT, RespBERT, MacBERT, DistilBERT	Applied in 22.45% (11/49) of studies ● ED symptom and history extraction [76] ● Emergency syncope differential diagnosis recognition [98] ● EM gout flare early prediction [81]
		Large Language Model Meta AI (LLaMA)	LLaMA is a family of open-source LLMs developed by Meta AI, built on the transformer architecture and trained on	LLaMA-2, LLaMA-3, LLaMA-7B, LLaMA-65B, PMC_LLaMA_7B, Vicuna-13B	Applied in 16.33% (8/49) of studies ● Pharmacotherapy recommendation evaluation [103] ● PICU differential diagnosis generation [64] ● Pediatric ESI prediction [99]

			publicly available text to support efficient and scalable language understanding [48,152].		- ED triage severity classification [79] - ED radiology report concept extraction [75]
		Mixtral of Experts (Mixtral)	Mixtral is a sparse mixture of expert language models based on the Transformer architecture, where each layer selects two out of eight feedforward expert blocks to process each token, allowing the model to use only a subset of its total parameters during inference [153].	Mistral-Large, BioMistral 7B	Applied in 4.08% (2/49) of studies Emergency heart attack risk stratification [89]
		Text-to-Text Transfer Transformer (T5)	T5 is a Transformer-based model that frames all NLP tasks into a unified text-to-text format, using the same model, objective, and training	T5-small, T5-base	Applied in 4.08% (2/49) of studies ED consultation note summarization [86]

			procedure across diverse tasks [154].		

Appendix eTable 4: Multimodal Foundation Models Relevant to Emergency and Critical Care

Overview of multimodal foundation model families relevant to ECC. This table presents representative multimodal FMs, such as CLIP, BLIP, and DALL·E, along with their architectural definitions. These models integrate cross-modal information but were not directly applied in ECC tasks within the included studies. They are included to provide structural context for the broader FM landscape.

Category	Category Definition	Representative Model Series	Model Series Definition
Multimodal Foundation Model	Multimodal model refers to FMs pretrained on large-scale cross-modal data using alignment and fusion mechanisms to integrate information from two or more modalities, enabling it to perform general understanding, reasoning, or generation tasks across them [155–158].	Contrastive Language–Image Pretraining (CLIP)	CLIP is a model consisting of an image encoder and a text encoder, jointly trained with contrastive learning to map images and text into a joint multimodal embedding space that aligns images with their corresponding text, enabling zero-shot performance on diverse vision–language tasks without task-specific fine-tuning, and leveraging large-scale image-text pairs with natural language supervision to achieve broad cross-task generalization [50].

		Bootstrapped Language-Image Pretraining (BLIP)	BLIP is a vision-language pre-training framework based on a Multimodal Mixture of Encoder-Decoder (MED) architecture that learns unified vision-language understanding and generation capabilities by jointly optimizing image-text contrastive learning, image-text matching, and image-conditioned language modeling, incorporating a Captioning and Filtering (CapFilt) dataset bootstrapping method to improve learning from large-scale noisy image-text pairs [51].
		Zero-Shot Text-to-Image Generation (DALL·E)	Zero-Shot Text-to-Image Generation (DALL·E) is a large-scale generative model based on a transformer that autoregressively models text and image tokens as a single stream of data, enabling zero-shot image generation controllable through natural language descriptions [159].

Appendix eTable 5: Vision and Audio Foundation Models Relevant to Emergency and Critical Care

Overview of vision and audio foundation models with potential relevance to emergency and critical care. This table describes foundational model families for image (e.g., Vision Transformers) and audio (e.g., Wav2Vec 2.0) data, summarizing their architectures and pretraining objectives. These vision and audio FMs were not directly used in ECC tasks in the reviewed studies, but are included to contextualize FM categories compatible with ECC data modalities.

Category	Representative Model Series	Model Series Definition
Vision Foundation Model	Vision Transformers (ViT)	Vision Transformer (ViT) is a model that applies a standard Transformer encoder directly to a sequence of linearly embedded image patches with learnable position embeddings, trained in a supervised fashion for image classification and achieving excellent results compared to state-of-the-art convolutional networks when pre-trained on large datasets [49].
Audio Foundation Model	Waveform-to-Vector 2.0 (Wav2Vec 2.0)	Wav2Vec 2.0 is a framework for self-supervised learning of speech representations that encodes raw audio with a convolutional feature encoder, builds contextualized representations with a Transformer, then trains by masking spans of the latent speech representations while solving a contrastive task over a product quantization of these representations before fine-tuning for speech recognition [160].