



Article  
scientifique

Revue de la  
littérature

2023

Published  
version

Open  
Access

This is the published version of the publication, made available in accordance with the publisher's policy.

---

## Patient information summarization in clinical settings : scoping review

---

Keszthelyi, Daniel; Gaudet-Blavignac, Christophe; Bjelogrljic, Mina; Lovis, Christian

### How to cite

KESZTHELYI, Daniel et al. Patient information summarization in clinical settings : scoping review. In: JMIR medical informatics, 2023, vol. 11, p. e44639. doi: 10.2196/44639

This publication URL: <https://archive-ouverte.unige.ch/unige:176647>

Publication DOI: [10.2196/44639](https://doi.org/10.2196/44639)

© The author(s). This work is licensed under a Creative Commons Attribution (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0>

Review

# Patient Information Summarization in Clinical Settings: Scoping Review

Daniel Keszthelyi<sup>1,2</sup>, MSc; Christophe Gaudet-Blavignac<sup>1,2</sup>, PhD; Mina Bjelogrić<sup>1,2</sup>, PhD; Christian Lovis<sup>1,2</sup>, MPH, MD

<sup>1</sup>Division of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland

<sup>2</sup>Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland

**Corresponding Author:**

Daniel Keszthelyi, MSc

Division of Medical Information Sciences, University Hospitals of Geneva

Rue Gabrielle-Perret-Gentil 4

Geneva, 1205

Switzerland

Phone: 41 223726201

Email: [Daniel.Keszthelyi@unige.ch](mailto:Daniel.Keszthelyi@unige.ch)

## Abstract

**Background:** Information overflow, a common problem in the present clinical environment, can be mitigated by summarizing clinical data. Although there are several solutions for clinical summarization, there is a lack of a complete overview of the research relevant to this field.

**Objective:** This study aims to identify state-of-the-art solutions for clinical summarization, to analyze their capabilities, and to identify their properties.

**Methods:** A scoping review of articles published between 2005 and 2022 was conducted. With a clinical focus, PubMed and Web of Science were queried to find an initial set of reports, later extended by articles found through a chain of citations. The included reports were analyzed to answer the questions of where, what, and how medical information is summarized; whether summarization conserves temporality, uncertainty, and medical pertinence; and how the propositions are evaluated and deployed. To answer how information is summarized, methods were compared through a new framework “collect—synthesize—communicate” referring to information gathering from data, its synthesis, and communication to the end user.

**Results:** Overall, 128 articles were included, representing various medical fields. Exclusively structured data were used as input in 46.1% (59/128) of papers, text in 41.4% (53/128) of articles, and both in 10.2% (13/128) of papers. Using the proposed framework, 42.2% (54/128) of the records contributed to information collection, 27.3% (35/128) contributed to information synthesis, and 46.1% (59/128) presented solutions for summary communication. Numerous summarization approaches have been presented, including extractive (n=13) and abstractive summarization (n=19); topic modeling (n=5); summary specification (n=11); concept and relation extraction (n=30); visual design considerations (n=59); and complete pipelines (n=7) using information extraction, synthesis, and communication. Graphical displays (n=53), short texts (n=41), static reports (n=7), and problem-oriented views (n=7) were the most common types in terms of summary communication. Although temporality and uncertainty information were usually not conserved in most studies (74/128, 57.8% and 113/128, 88.3%, respectively), some studies presented solutions to treat this information. Overall, 115 (89.8%) articles showed results of an evaluation, and methods included evaluations with human participants (median 15, IQR 24 participants): measurements in experiments with human participants (n=31), real situations (n=8), and usability studies (n=28). Methods without human involvement included intrinsic evaluation (n=24), performance on a proxy (n=10), or domain-specific tasks (n=11). Overall, 11 (8.6%) reports described a system deployed in clinical settings.

**Conclusions:** The scientific literature contains many propositions for summarizing patient information but reports very few comparisons of these proposals. This work proposes to compare these algorithms through how they conserve essential aspects of clinical information and through the “collect—synthesize—communicate” framework. We found that current propositions usually address these 3 steps only partially. Moreover, they conserve and use temporality, uncertainty, and pertinent medical aspects to varying extents, and solutions are often preliminary.

(JMIR Med Inform 2023;11:e44639) doi: [10.2196/44639](https://doi.org/10.2196/44639)

**KEYWORDS**

summarization; electronic health records; EHR; medical record; visualization; dashboard; natural language processing

## Introduction

### Background

Summarization is an essential element of human cognition and consists of taking a set of information and retaining the pertinent elements to take action [1]. Feblowitz et al [2] defined information summarization in the health care context as “the act of collecting, distilling, and synthesizing patient information for the purpose of facilitating any of a wide range of clinical tasks.” This definition translates as simplifying the presented information so that health care professionals (HCPs) can act more smoothly and efficiently in different clinical situations.

Automatic summarization of information in electronic health records (EHRs) can serve as a solution for information overload [3], a widespread problem in health care when the presented data are too much to be efficiently processed in a care situation. Information overload can have detrimental effects on patient care in the form of professional stress, fatigue, delays, and medical errors [4]. Although the phenomenon is not novel, it is increasingly present owing to an aging population with an exponentially increasing presence of chronic diseases, increased administrative burden, overabundance, and suboptimal storage of medical data [5,6]. Furthermore, current EHR systems present information in a fragmented manner [7] with widespread repetition, copy-pasting [8], and details not relevant to clinical care [9].

Despite the need for automatic patient information summarization, there is no widely accepted theory or methodology. This report aimed to synthesize the contributions of patient information summarization scattered in the literature. Scoping review is the chosen form with the aim of mapping ideas, mapping concepts related to the question, and identifying knowledge gaps.

The review is not unprecedented: in their narrative review, Pivovarov and Elhadad [10] already summarized the most important contributions to clinical summarization in 2015. Moreover, there have been several published studies surveying the literature in related fields, including the summarization of biomedical literature [11,12], the summarization from medical documents [13], neural natural language processing (NLP) in EHRs [14], named entity recognition, a type of information extraction and NLP technique, free-text clinical notes [15], automatic clinical documentation [16], the visualization of medical information in the clinical context [17-20], the visualization of intensive care unit (ICU) data [21], and the visualization of trends in medical data [22]. The latter reviews, although exhaustive in their specific scope, do not permit the identification of state-of-the-art summarization methods for HCPs. For example, it is difficult to state the current state of research for the management of uncertainty and time in clinical summarization. Moreover, they did not provide any informed guidelines for clinical summarization.

### Objective

This review, building on a broader scope of articles than the combination of all the previous studies, systematically evaluates where, what, and how medical information is summarized; whether summarization conserves temporality, uncertainty, and medical pertinence; and how the propositions are evaluated and deployed.

On the basis of cognitive science literature, this review also proposes a novel “collect—synthesize—communicate” framework to compare studies on how they contribute to clinical summarization.

## Methods

### Overview

The methodology was designed to process a broad scope of articles; hence, different search strategies were combined to diversify the sources. Two reviewers agreed on the selection method: 2 databases with a clinical focus were searched with similar queries and the retrieved articles were filtered by one of the reviewers according to their titles and abstracts. The same reviewer read the remaining reports in the full text and selected them according to the inclusion and exclusion criteria. The same filtering was then carried out on citations within these articles and the citations of these articles. The reporting was done using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [23], and a checklist is provided in [Multimedia Appendix 1](#).

The 2 databases searched in this review were the Web of Science Core Collection and PubMed as they contain a broad scope of articles in the medical field and are less inclusive of other articles in computer science, not related to the medical or scientific domain.

The search query for Web of Science was designed as a combination of 2 parts: capturing the summarization process and capturing the health care content. An iterative process was used to define the exact search term, where the gain of adding a keyword was examined by determining whether the first 5% (sorted by relevance defined by Web of Science) of the results from a query containing the new word but excluding the previous words shows any relevant article. This led to the following query: “ALL=((‘ehr’ OR ‘emr’ OR ‘health’ OR ‘patient’ OR ‘medical’ OR ‘hospital’ OR ‘healthcare’) AND (‘summarization’ OR ‘summarisation’ OR ‘summarizing’))” searching in the title, abstract, and metadata of the articles in the database, including the “keyword plus” field containing terms frequently appearing in the body of an article but not mentioned in the title or abstract.

The query in PubMed was “(‘her’ OR ‘emr’ OR ‘health’ OR ‘patient’ OR ‘medical’ OR ‘hospital’ OR ‘health care’ OR ‘medical record’[(MeSH Terms)]) AND (‘summarization’ OR ‘summarisation’ OR ‘summarizing’).” This query was almost identical to the search made in Web of Science except that

PubMed does not have a “keyword plus” field like Web of Science to search in. Instead, preindexed articles with the Medical Subject Headings term “medical records” were included in the search.

All results of the queries were imported into the Rayyan app [24], helping to organize the citations for a review article. After duplicates were removed, the abstracts and titles were scanned in this application to filter records that were obviously irrelevant to patient information summarization. The app enables highlighting specific words in the abstract with 2 distinct colors, speeding up the review process. After this filtering step, the remaining articles were read in the full text for inclusion using the inclusion and exclusion criteria detailed in the following section.

After identifying the relevant works, the list of references in the selected articles and the list of citing papers retrieved by Google Scholar were manually reviewed for titles related to the topic. These potentially relevant titles were manually added to the citation manager and further filtered by reading their abstract and eventually reading them in the full text (as with the original results). If these articles contained further (previously unseen) relevant references or citations toward them, they were also processed.

### Inclusion and Exclusion Criteria

According to the inclusion criteria, all records mentioning the summarization of clinical or health data as a general goal in their abstract, proposing solutions for information overflow, or claiming to make steps for these general goals were included.

All records were excluded where the corresponding article was not available in the full text for the authors (ie, at the University of Geneva campus), as the analysis could not be conducted on these records.

As several articles do not mention the source of the data in their abstracts, an exclusion criterion excludes articles about summarizing non-EHR data (eg, summarizing research articles; not EHR). Similarly, articles developing summarization for

users other than HCPs (*not for HCP*) or for contexts other than clinical applications (*not clinical*) were excluded.

As many different methods can be labeled as “summarization,” only records presenting a type of overview of a patient’s current or past status are aimed to be included, therefore articles proposing alerts (eg, risk scores and cues) or similar simple parameters to summarize the state of a disease or patient (*alert*) and articles proposing other remedies for information overflow than automatic summarization for information overflow (*not automatic*) were excluded.

As previous reviews analyzed articles using different aspects, a broad timeframe was aimed at the review. However, as early EHR systems were very different from current systems, and hence the concept of summarization is largely different in these systems, articles before 2005 were excluded (<2005). The cutoff year is somewhat of an arbitrary (but round) threshold, although contributions before this year are sporadic.

Finally, articles presenting summarization solutions only for nontextual and nonstructured data (eg, video or signal summarization; *Other data*) and review papers (*Review*) were also excluded.

Articles found relevant to the review were evaluated by one of the reviewers for several criteria chosen to answer the following questions:

1. Where is summarization performed?
2. What is summarized? How?
3. How crucial aspects of clinical information are conserved and used?
4. How are the algorithms evaluated?

**Textbox 1** presents the detailed criteria. Some of these criteria were defined a priori, whereas others were shaped during the analysis process. For 1 aspect, the input data type for summarization, the analysis was carried out on a broader scope, and reports excluded by the “other data” criterion were also analyzed for this information.

**Textbox 1.** Criteria according to which articles are evaluated in the analysis part of this review. For some of the criteria, categories are defined a priori. For others, they are shaped during the analysis process (the ones defined a posteriori are marked with an asterisk).

#### General aspects

- Type of the study
  - Prototype: articles describing a summarization system or algorithm that can be evaluated. The evaluation might be present or absent from the article.
  - Evaluation study: articles evaluating summarization systems, algorithms, or current summarization processes in health care without presenting a new automatic summarization solution.
  - Recommendations: articles with theoretical contributions not being implemented.

#### Where are summaries needed?

- Field of application\*: the medical or clinical domain where the summarization is applied. The categories are discovered during the review process.

#### What should be summarized?

- Source of information:
  - Single encounter
  - Multiple encounters
  - This information cannot be inferred from the text.
- Input for the summarization:
  - Structural data: a combination of numerical and categorical data
  - Textual data: free-text patient information present in electronic health record systems

#### How to summarize?

- The summarization method\*: the categories of summarization methods are shaped during the review process.
- Presentation\*: how the summary is presented to the end user. The types of presentations are shaped during the review process.
- View on the summarization problem:
  - The top-down group represents records where summarization consists of eliminating “disturbances” from all available information, that is, hiding information deemed to be unnecessary.
  - Bottom-up methods see summarization as a process of finding the most salient information available and building a summary from it.

#### Aspects to be conserved during summarization

- Temporality\*: if and how temporal information is conserved and used during summarization. The categories are shaped by the discoveries in the scoping review.
- Uncertainty\*: if and how the uncertainty of information is represented during summarization. The categories are shaped by the discoveries in the scoping review.
- Medical knowledge\*: if any medical knowledge is included in the design of the summarization system or during summarization. The categories are shaped during the review process.

#### What is a good summary? Evaluation and deployment

- Evaluation\*: the method of evaluation. The types are shaped according to the discoveries from the review process.
- Deployment: if the summarization system was deployed in real clinical settings

### Collect—Synthesize—Communicate Framework

During the analysis, we developed a new framework to compare methods of how they summarize clinical information. Following the definition presented in the introduction [2], the model divides the summarization process into an ideally sequential process of information collection, information synthesis, and summary communication. Information collection refers to the extraction

of information from raw data, synthesis describes the selection and eventual transformation of the retrieved information, and communication refers to the representation of the synthesized information in a human-digestible format.

This view was consistent with that of several sources of cognitive psychology. For example, Johnson [25] describes summarization as a sequence of prerequisites for summarization (including comprehending individual propositions of a story,



establishing connections, identifying the consistent structure of the story, and remembering the information), information selection, and formulating a concise summary. This is also similar to the view presented by Hidi and Anderson [26], who discussed selection, condensation, and transformation.

Nevertheless, the few theoretical studies on clinical summarization have slightly different views on the process. Feblowitz et al [2] described clinical summarization as a process

of aggregation, organization, reduction, transformation, interpretation, and synthesis. Jones [27,28] describes textual summarization as a process of interpretation, transformation, and text generation. Although these theories mention seemingly different steps for summary creation, they can be mapped to the proposed simpler and more general 3-step framework. Table 1 presents the mapping to the present framework of these theories and some of the most commonly used summarization methods.

**Table 1.** Summary of how existing theoretical frameworks and most abundant summarization methods relate to the collect—synthesize—communicate framework.

Theory or method	Collection	Synthesis	Communication
Feblowitz et al [2]	Aggregation, organization, and interpretation	Reduction, transformation, and synthesis	Organization and synthesis
Jones [27,28]	Interpretation	Transformation	Text generation
Extractive summarization (eg, Liang et al [29])	N/A <sup>a</sup> (not covered)	Sentence selection	N/A (not covered)
Abstractive summarization (eg, Gundogdu et al [30])	Encoding	Attention mechanism	N/A (not covered)
Topic modeling (eg, Botsis et al [5])	Topic extraction	N/A (not covered)	N/A (not covered)

<sup>a</sup>N/A: not applicable.

Results

Overview

As shown in Multimedia Appendix 2 [31], a total of 7925 titles were retrieved from PubMed and 3641 articles were retrieved from Web of Science. After removing duplicates, 9166 records were screened by their title and abstract for inclusion criteria and 380 records were chosen for full-text reading. From these, 1 could not be accessed by the authors and 328 were excluded based on the exclusion criteria.

From the 52 articles included in the analysis, 612 records were identified as potentially relevant by their title and 175 titles were chosen to be read in the full text after screening the abstracts. From these 175 titles, 2 could not be accessed, 97 records were excluded according to the exclusion criteria, and 76 titles were included in the analysis.

Among the 128 articles remaining in the analysis, 102 titles were categorized as a prototype, 20 were categorized as evaluation studies, and 6 were categorized as “recommendations.”

Fields of Application

This review identified diverse fields of application for which summarization methods have been developed. A grouping of these fields is as follows:

1. ICU (27/128, 21.1%), where recent events and vital parameters are summarized
2. *Surgery* (1/128, 0.8%) and *related anesthesiology* (5/128, 3.9%), requiring all the information related to surgery to be summarized
3. *Diagnostics*, showing findings from one or several diagnostic sessions and including radiology (19/128, 14.8%), out of which 5.5% (7/128) were presented as a solution in the MEDIQA 2021 summarization task [32],

- ultrasound (2/128, 1.6%), prostatectomy (1/128, 0.8%), and laboratory data management in a clinical context (1/128, 0.8%)
4. Hospital care (9/128, 7%), where information related to a hospital stay requires efficient summarization
5. Chronic disease monitoring including diabetes (4/128, 3.1%), HIV (1/128, 0.8%), chronic obstructive pulmonary disease care (1/128, 0.8%), cardiology (2/128, 1.6%), nephrology (1/128, 0.8%), and monitoring of multiple chronic diseases (4/128, 3.1%), where salient events and information during a complex and long-lasting disease are required
6. Oncology (5/128, 3.9%), where the main events and elements of complex treatment are summarized
7. Drug prescription (3/128, 2.3%), where pharmaceutical history is summarized
8. Other medical environments included psychotherapy (3/128, 2.3%), opioid misuse treatment (1/128, 0.8%), general practice (2/128, 1.6%), emergency room (2/128, 1.6%), older adult care (2/128, 1.6%), and maternal care (1/128, 0.8%).

In addition, 25% (32/128) of articles did not specify their field of application or were meant to be usable in multiple types of medical environments and domains.

Input for Summarization

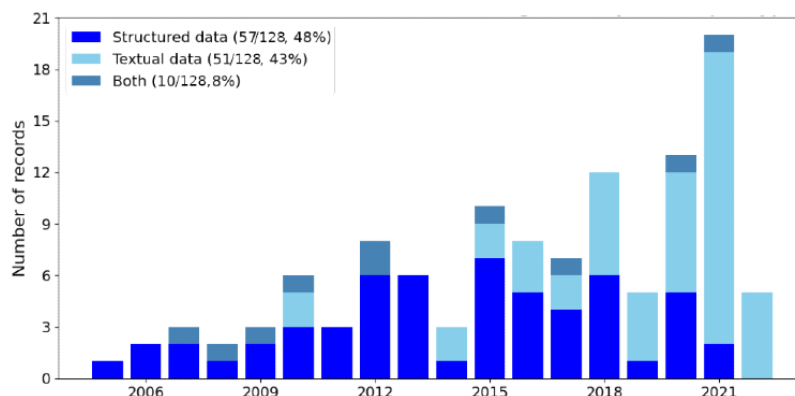
Regarding the source of information, 62.5% (80/128) of reports talk about systems summarizing single patient encounters, 27.3% (35/128) of reports explicitly talk about summarizing multiple encounters, 6.3% (8/128) of reports implicitly describe multiple encounter summarization, and 3.9% (5/128) of reports did not specify the cardinality of encounters.

Among the 128 articles in the review, 3 (2.3%) reports *do not specify the input* type for the summary, 59 (46.1%) worked only with *structured data*, 53 (41.4%) worked only with *textual data*,

and 13 (10.2%) worked with *both types*. The trends in the number of articles with different input types are shown in Figure 1. Although more records use only structured data as the input type, the number of articles using textual information has shown

a rapidly increasing trend in recent years. Textual information is usually assumed to be in English [33,34] and presents solutions for Finnish and German languages [35,36].

**Figure 1.** The number of records by year of publication and the input type used in the summarization system or method presented or evaluated. Each column corresponds to a year, the different input types are aggregated into this column, their proportion for the given year is visible on the figure. ICU: intensive care unit.



A rapid analysis of the excluded articles using other types of clinical data identified the following:

- Overall, 37 video and image sequence visualization tools in the clinical domain, mainly keyframe extraction [37,38] or motif discovery [39] methods in various fields of medicine, including older adult care, endoscopy [40], hysteroscopy [40,41], laparoscopy [42], magnetic resonance image [39,43], and ultrasound [44]
- Overall, 20 sensory data simplification techniques using time-series analysis, motif discovery [45], and classification [46] methods for electrocardiogram or other types of signals
- Overall, 117 articles about summarizing genomic data [47]

## How Data Can Be Summarized? Summarization Methods

Common summarization methods used in the analyzed studies include (a report might use several of these) the following:

- Visual design* (59/128, 46.1%) organizes the information visually to help HCPs understand it within a short timeframe.
- Concept and relation extraction* (30/128, 23.4%): extracts semantic information from textual information
- Abstractive summarization* (19/128, 14.8%) [30,48-52] shortens texts by reformulating them using different wording to describe the content of a document [10].
- Extractive summarization* (13/128, 10.2%) [29,53] shortens texts by omitting a part of it, that is, composing a short text (a summary) from extracts of the original document.
- Summary specification* (11/128, 8.6%) describes the content to be presented for a summary.
- Pipeline extracting information and synthesizing and communicating* it with natural language generation tools (7/128, 5.5%)
- Topic modeling* (5/128, 3.9%) [54-57] categorizes documents according to their content and labels them with a list of representative words [58]
- Time-series analysis* (6/128, 4.7%) identifies characteristic properties in a temporal data series, including motif

discovery, identifying meaningful patterns in temporal data (used in the study by Jane et al [59]), trend detection [60], or change detection [61].

- Dimensionality reduction* (3/128, 2.3%) treats patient data as a long vector encoding all patient information (ie, a row in a table with many columns), and reduces this information to a shorter vector (ie, a row with a much smaller number of columns) without losing too much information.

Some of these methods are intrinsic to the input data type and work only with a particular data type. For example, time-series analysis (including motif discovery), risk scores, and dimensionality reduction are intrinsic methods for structured data. Although a large number of articles using these methodologies are not included in this review as they are used by non-HCPs (eg, machine learning algorithms), some of the titles propose this approach as the first step to clinical summarization [62-64].

The most common intrinsic methods for textual data are extraction, abstractive summarization, and topic modeling.

Some of these summarization methods can apply machine learning techniques. An overview of the applied machine learning methods is presented in Table 2. The table lists all records obtained using machine learning and categorizes the records according to the summarization method and the type of machine learning method they use. Machine learning methods can be categorized into traditional machine learning methods, deep neural networks, and transformers. Traditional methods include support vector machines, random forests, and conditional random field methods; deep neural networks include deep neural networks, recurrent neural networks, and convolutional neural networks; transformers contain BART [65], BERT [66], Pegasus-based [67] methods, and pointer-generator models. In addition, Reunamo et al [34] used an interpretable machine learning technique (Local Interpretable Model-Agnostic Explanations, LIME [68]). N/A indicates that a machine learning method is not used for a given type of summarization method.

**Table 2.** Summary of records applying machine learning methods for clinical summarization<sup>a</sup>.

	Traditional techniques	Deep neural networks	Transformers
Extractive summarization	<ul style="list-style-type: none"> <li>SVM<sup>b</sup>+CRF<sup>c</sup> <ul style="list-style-type: none"> <li>Liang et al [29], 2019</li> <li>Liang et al [9], 2021</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>CNN<sup>d</sup> <ul style="list-style-type: none"> <li>Liang et al [29], 2019</li> <li>Liang et al [9], 2021</li> <li>Subramanian et al [69], 2021</li> </ul> </li> <li>RNN<sup>e</sup> <ul style="list-style-type: none"> <li>Alsentzer and Kim [70], 2018</li> <li>Liu et al [71], 2018</li> <li>Chen et al [53], 2019</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>BERT<sup>f</sup> <ul style="list-style-type: none"> <li>Chen et al [53], 2019</li> <li>Kanwal and Rizzo [72], 2022</li> <li>McInerney et al [73], 2020</li> <li>Liang et al [36], 2022</li> <li>Shah and Mohammed [56], 2020</li> </ul> </li> <li>Other <ul style="list-style-type: none"> <li>Liang et al [29], 2019</li> <li>Liang et al [9], 2021</li> </ul> </li> </ul>
Abstractive summarization	<ul style="list-style-type: none"> <li>N/A<sup>g</sup></li> </ul>	<ul style="list-style-type: none"> <li>RNN <ul style="list-style-type: none"> <li>Gundogdu et al [30], 2021</li> <li>Hu et al [74,75], 2021</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>BERT <ul style="list-style-type: none"> <li>Cai et al [48], 2021</li> <li>Chang et al [76], 2021</li> <li>Mahajan et al [77], 2021</li> <li>Sotudeh et al [50], 2020</li> </ul> </li> <li>BART<sup>h</sup> <ul style="list-style-type: none"> <li>Dai et al [78], 2021</li> <li>He et al [79], 2021</li> <li>Kondadadi et al [80], 2021</li> <li>Shing et al [81], 2021</li> <li>Xu et al [82], 2021</li> </ul> </li> <li>Pegasus <ul style="list-style-type: none"> <li>Dai et al [78], 2021</li> <li>He et al [79], 2021</li> <li>Kondadadi et al [80], 2021</li> <li>Zhu et al [89], 2021</li> <li>Xu et al [82], 2021</li> </ul> </li> <li>Pointer generator <ul style="list-style-type: none"> <li>MacAvaney et al [49], 2019</li> <li>Zhang et al [51], 2018</li> <li>Zhang et al [83], 2019</li> </ul> </li> <li>Own architecture <ul style="list-style-type: none"> <li>Delbrouck et al [84], 2021</li> </ul> </li> <li>GPT-2<sup>i</sup> <ul style="list-style-type: none"> <li>Xu et al [85], 2019</li> </ul> </li> </ul>
Concept and relation extraction	<ul style="list-style-type: none"> <li>N/A</li> </ul>	<ul style="list-style-type: none"> <li>RNN: <ul style="list-style-type: none"> <li>Reunamo et al [34], 2022</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>BART: <ul style="list-style-type: none"> <li>Tang et al [86], 2022</li> </ul> </li> </ul>
Pipeline	<ul style="list-style-type: none"> <li>Random forest: <ul style="list-style-type: none"> <li>Lee and Uppal [87], 2020</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>N/A</li> </ul>	<ul style="list-style-type: none"> <li>N/A</li> </ul>
Topic modeling	<ul style="list-style-type: none"> <li>Alternating decision tree: <ul style="list-style-type: none"> <li>Devarakonda et al [88], 2017</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>N/A</li> </ul>	<ul style="list-style-type: none"> <li>N/A</li> </ul>

<sup>a</sup>Machine learning methods are categorized into traditional machine learning methods, deep neural networks, and transformers.<sup>b</sup>SVM: support vector machine.<sup>c</sup>CRF: conditional random field.<sup>d</sup>CNN: convolutional neural network.<sup>e</sup>RNN: recurrent neural network.<sup>f</sup>BERT: Bidirectional Encoder Representation from Transformers.<sup>g</sup>N/A: not applicable.<sup>h</sup>BART: Bidirectional Autoregressive Transformer.<sup>i</sup>GPT-2: Generative Pre-trained Transformer 2.

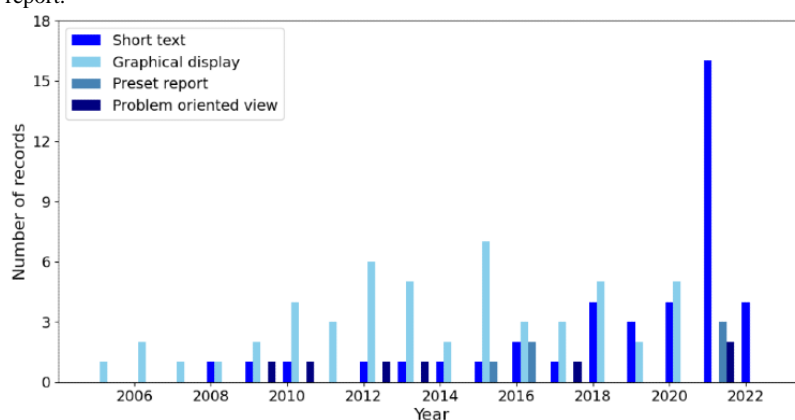


Summarization methods can also be categorized based on their outputs. The review identified several ways in which the summarized information is presented to the end user:

- A graphical display (53/128, 41.4%) is a specific way (interactive or not) to present information on the computer screen.
- A short textual summary (41/128, 32%) describes information in an ordinary language (eg, English).
- A preset static report: including its content designed to include specific medical information (6/128, 4.7%) or chosen statistical distributions representative of the patient (1/128, 0.8%)
- Problem-oriented view: a view grouping findings according to the problems the patient may present (7/128, 5.5%).
- Low-dimensional vector (4/128, 3.1%): encoding information in numbers (where  $n$  is the dimension) where each number represents the state of the patient from a particular aspect.
- List of words representing a topic (5/128, 3.9%), problem list (2/128, 1.6%), list of medical concepts found in the document (2/128, 1.6%), or label (2/128, 1.6%)
- A table (1/128, 0.8%) with rows and columns or a directed graph or concept map representing information in a graph-structured data model (2/128, 1.6%)
- No presentation: the articles in the “recommendation” group (5/128, 3.9%) did not present the results to the end user.

Figure 2 depicts the evolution (by the time of publication of records) of the most abundant formats for communicating the summarization results.

**Figure 2.** The number of records by year of publication and the most common ways of summary presentation used in the summarization method presented or evaluated in the report.



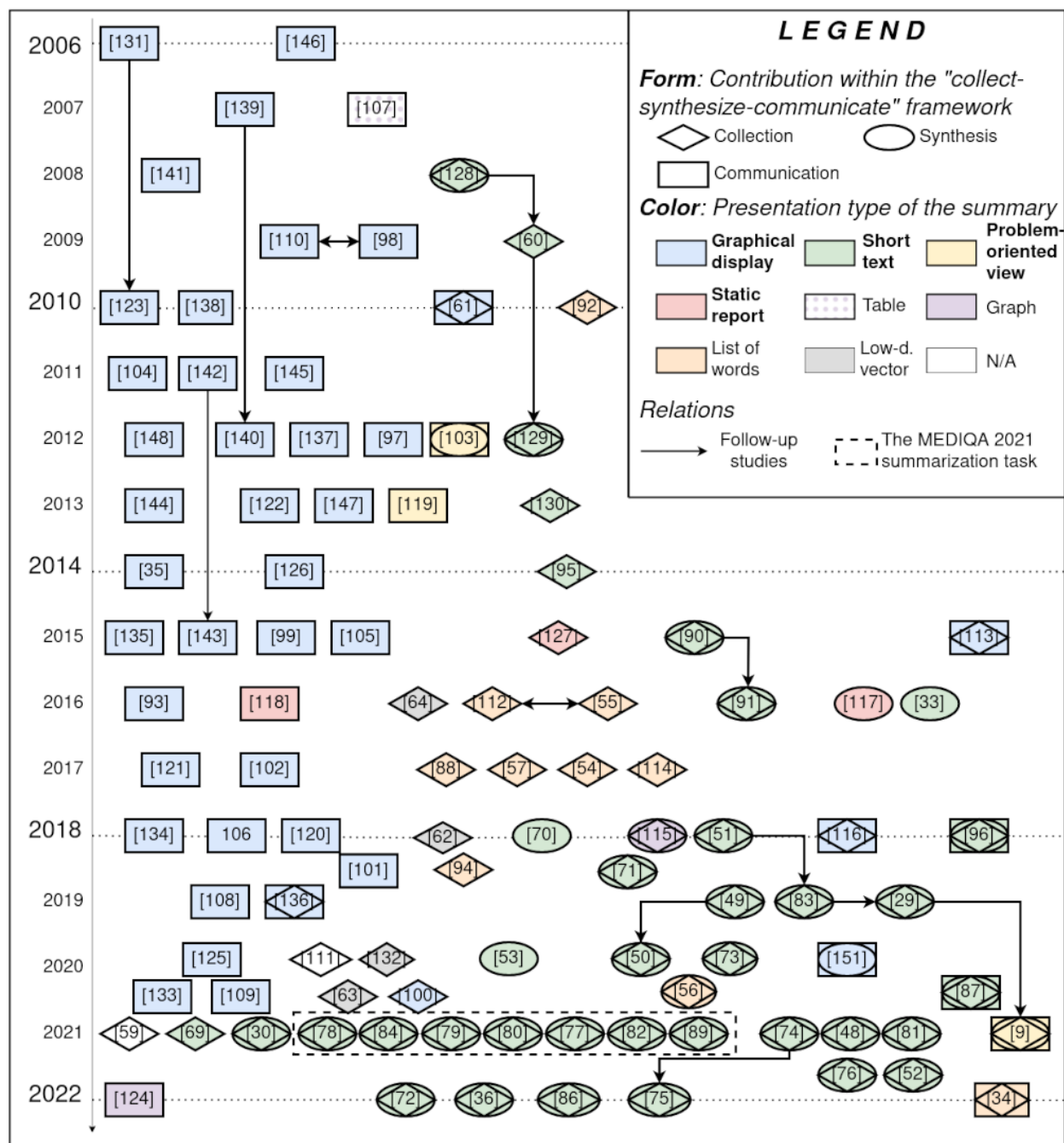
Concerning their view on summarization, 32.8% (42/128) of the records regarded summarization as a *bottom-up* approach and 64.8% (83/128) used the *top-down* view, whereas 1.6% (2/128) of records *do not show a clear opinion* on summarization.

Using the proposed framework, 42.2% (54/128) of the records contributed to *information collection*, 33.6% (43/128) recorded *information synthesis*, and 46.1% (59/128) presented solutions for *summary communication*. Figure 3 [9,29,30,33-36,48-57,59-64,69-84,86-149] visualizes all the analyzed prototype articles and how they fall into these categories (ie, which step of the framework is addressed within the corresponding work). The records' year of publication, presentation of summaries, and relationship between records

are also displayed. The diagram has a vertical axis showing the year of publication, and all the “prototype” records (presented as a reference) published in that year appear in a line (or in 2 lines if the number of publications for a given year is very high). The order within a line has no significance, although the records were grouped within a line to show their contributions. The shape or shapes surrounding a reference symbolizes the steps that the record addresses in the “collect—synthesize—communicate” framework. The reference to the study by Liang et al [9] is surrounded by all 3 shapes, indicating that the study addresses all the 3 steps. The records also have a color representing in which format the summaries are presented to the HCPs. Closer relationships (ie, follow-up studies) are also presented. The studies submitted to the MEDIQA-2021 challenge [21] are also marked in the diagram.

**Figure 3.** Diagram showing references to all analyzed “prototype” records and how they contribute to the “collect—synthesize—communicate” framework (ie, which step of the framework is addressed within the corresponding work). The records’ year of publication, the presentation of summaries, and the relation between records are also displayed [9,29,30,33-36,48-57,59-64,69-84,86-149].

**Overview of the “prototype” records,  
their contribution to the proposed framework,  
and their types of summary presentation**



Regarding information collection, concept and relation extraction (50/54, 56%), time-series analysis (6/54, 11%), encoding (22/54, 41%), temporal abstraction (6/54, 11%), and topic extraction (5/54, 9%) were proposed as solutions. Medical concepts are extracted from textual data either using publicly available solutions (eg, cTAKES [164] in the study by Goff and Loehfelm [94]) or tools developed by the authors (eg, [113,114,157]). The retrieved list of concepts can be used for simpler tasks, such as problem list generation [88], or some

records present systems that take a step further extracting the context [115], syntactic structure [94], or approximate semantic structure of a sentence [116] as well.

With regard to information synthesis, sentence selection by scoring (13/43, 30%), knowledge-based rules (18/43, 42%), and attention mechanism (19/43, 44%) were possible solutions.

Proposals for summary visualizations are usually features on a graphical screen; they are listed and compared in Table 3. For

unprocessed textual data, the solutions included highlighting important concepts (3/5, 60%) and creating graphs that visualize the semantic structure of the textual data (2/5, 40%).

**Table 3.** The number of records presenting various features for visualizations in works with graphical displays. A record can use several features.

Feature	Occurrence (out of 58), n
Colors	43
Selection of features	37
Tabular interface	35
Change in time	31
Visualization of divergence	22
Placement of variables	19
Interactive display	18
Pictograms	10
Physical location shown	6
Alerts	5
Size difference	5
Customizability	3
Shape	3
Word cloud	3
Comparison	2
Variability of parameters	1

### Aspects to Be Conserved and Used

In total, 58.6% (75/128) of the titles did not conserve temporal information, whereas 2.3% (3/128) of titles were agnostic to temporal information (they conserve but do not use it). The remaining articles used a variety of approaches:

- *Timeline visualization*: plotting information along a horizontal or vertical temporal axis (34/128, 26.6%).
- *Other visualizations*: showing only the trend of parameters (1/128, 0.8%) or providing a complex visualization framework in which temporal information can be displayed and analyzed (1/128, 0.8%).
- *Information extraction from the temporal domain* by analyzing how the parameters change during the patient journey. This group included a time-series analysis (6/128, 4.7%), pattern recognition (2/128, 1.6%), and change detection (1/128, 0.8%). Time-series analysis (applied in several studies) [59,60,125-127,150] extracts statistical information from the temporal evolution of one or several variables. Pattern recognition [117,128] attempts to identify meaningful patterns in temporal data. Change detection [61] seeks important events that manifest in the trends and patterns of temporal variables. Some studies have revealed the relationship between these events.
- *A theoretical model of temporal events*, which can either describe more complex interactions between temporal events (6/128, 4.7%) or be very simple (eg, creating an order: 1/128, 0.8% or describing events with a single time [n=1]).

It is worth noting that timeline visualization was applied in 3 articles in the temporal information extraction and in 1 article in the complex model of the temporality group as well.

Regarding information uncertainty, 89.1% (114/128) of the articles did not consider the uncertainty of information. Others have proposed the following solutions:

- Statistical methods were used to treat *uncertainty in data*. These methods included correcting detectable errors (3/128, 2.3% [60,79,150]) and optimizing the statistical description of the data using robust statistics (1/128, 0.8% [62]).
- *Uncertainty of temporal events* was described (2/128, 1.6%).
- *Uncertainty of statements* was described by assigning them to uncertainty categories (3/128, 2.3% [71,74,75]) or using existing ontology (3/128, 2.3%).

Medical pertinence was not conserved in 34.4%, (44/128) of the studies (ie, they had no requirements that the summary had any relation to medical concepts or knowledge). A total of 35.9% (46/128) of records used medical knowledge to specify the information to be included in the summary and with what design. Other propositions included the following:

- Using ontologies to find and relate concepts within textual notes (20/128, 15.6% [87,92,93]), the use of Unified Medical Language System (UMLS) extraction tools (6/128, 4.7%) to extract them (eg, [94-96]), or improving the performance of abstractive summarization (2/128, 1.6% [76])
- Use of risk scores to create visualizations (3/128, 2.3% [97-99]) or the application of guidelines to assess risks (2/128, 1.6% [100,101])

- The use of medically salient rules to constrain summarization (3/128, 2.3% [29,61,102])
- Evaluation of the factual correctness of the created summaries integrated into reinforcement learning (2/128, 1.6% [79,83])
- Application of medical knowledge to select pertinent information (2/128, 1.6% [103,151])
- The use of medical knowledge to construct evaluation metrics (2/128, 1.6% [81,152])

## What Is a Good Summary? Evaluation and Deployment

Several types of evaluation methods and metrics are presented in the publications:

- Quantitative measurements in experiments with human participants (31/128, 24.2%)
- Quantitative measurements when summarization was performed in a real clinical environment (8/128, 6.3%)
- Interviews (7/128, 5.5%), focus groups (2/128, 1.6%), or surveys (19/128, 14.8%) asking the opinions of the users after exposure to the summarization system
- Intrinsic evaluation (25/128, 19.5%) of measuring quality by comparing the results to a ground truth
- Performance on a proxy task (ie, disease prediction; 10/128, 7.8%)
- Performance in identifying human-annotated concepts (9/128, 7%) or topics (2/128, 1.6%)

The distribution of the number of human evaluators is shown in [Multimedia Appendix 3](#). Two records [119,173] used significantly more evaluators than other solutions, which are represented as 2 distinct groups at the tail of the histogram. One of these records [173] is a large-scale survey, whereas the other [119] is a pilot study measuring user performance.

Although some records present several evaluation techniques, in 9.4% (12/128) of the articles, no evaluation is presented; in 1.6% (2/128) of articles, the evaluation is not detailed; and in 4.7% (6/128) of records, the evaluation consists of a subjective evaluation carried out by the authors of the article.

The metrics used in the evaluations are as follows:

- *Performance metrics* (eg, precision, recall, and *F* score) on a *prediction/classification task* measuring the “goodness” (validity) of predictions or classifications (used both in usability experiments and formative evaluations; 11/128, 8.6%)
- *Performance metrics* (eg, accuracy) of human participants (ie, the validity of their decisions) on an *experimental task* (24/128, 18.8%)
- *Time savings* due to summarization systems: *time to completion* (ie, the time needed to perform a predefined task) in experiments (21/128, 16.4%) or time saved during patient visits (1/128, 0.8%) in deployed systems
- *Patient outcome metrics* (6/128, 4.7%) included mortality and hospital readmission rates.
- The *NASA-TLX score* describes the workload of the user (5/128, 3.9%) and the relationship between the NASA-TLX score and error count (1/128, 0.8%).

- *Number of interactions* (eg, click and screen change) in usability studies (3/128, 2.3%)
- *Grades* given by human evaluators measuring the utility and usability of a system (13/128, 10.2%) or trust in it (1/128, 0.8%).
- Scores comparing textual summaries with properties of the input text. These scores included *Recall-Orientated Understudy for Gisting Evaluation (ROUGE)* [153] (20/128, 15.6%), *bilingual evaluation understudy* [154] (2/128, 1.6%), and *comparison between input and output distributions* (2/128, 1.6%).
- *Other properties of the output textual summaries* including *readability/fluency* (10/128, 7.8%), *accuracy or factual correctness* (5/128, 3.9%), *completeness* (7/128, 5.5%), and *overall quality* (7/128, 5.5%) in qualitative evaluations of textual outputs. Two (N=128, 1.6%) records distinguished between ontological and nonontological correctness.
- *Proxy measures for the faithfulness* of textual summarization (6/128, 4.7%)
- *Heuristics* derived from requirement specification (5/128, 3.9%)

The evaluation metrics used in quantitative evaluations usually depend on the method of summarization; for dimensionality reduction, it is often a performance metric to predict diseases; for extractive and abstractive summarizations, the ROUGE score [153] is the most commonly used metric, as it is considered the most reliable [32], and for topic modeling, it is its empirical likelihood [174].

For text summarization, evaluations with human participants are often carried out by annotators subjectively grading each produced summary along some metrics, including readability, factual correctness, and completeness. For other summarization methods, this task is usually approximated by either usability tests [134-139] or experiments [140-147] where performance and workload are measured. The few systems deployed in clinical settings are often evaluated by measuring patient outcomes or clinical indicators.

Reviewing the results of each report, some records compared the results with summarization methods in the general domain [48], and 6 (5%) [30,32,50,75,78,144] presented a comparison of clinical summarization methods. The distribution of cross-citations between articles, that is, the number of other publications appearing in the review cited by each report, is represented in [Multimedia Appendix 4](#). Furthermore, 80% of the records cited fewer than 3 other articles analyzed in this review.

Among the 128 records analyzed, 4 (3.1%) talked about a method deployed on a large scale, 7 (5.5%) described deployment in a pilot study, and 1 (0.7%) disclosed the code alongside the publication.



## Discussion

### Principal Findings

#### *Where Are Summaries Needed in Health Care?*

Publications on clinical summarization are tied to several different medical and clinical fields, mainly where quick decision-making is crucial (eg, ICU) or where a large amount of information is routinely produced (eg, oncology, chronic disease management, and hospital care).

However, some fields requiring quick decision-making (eg, emergency room environments) have seen less progress. In contrast, others where quick decision-making is less critical (eg, radiology) are covered by a relatively large number of records. This discrepancy suggests that clinical summarization can be beneficial in almost all medical fields, although the idea may not have reached all domains at the same pace. Although the previous drivers are easily identifiable, we speculate that the presence of other solutions proposed to handle information overload (eg, the study by Xu et al [85]; see the study by Hall and Walton [4] for review) can decelerate, whereas a shortage of personnel in a field (eg, radiology [155]) can accelerate adaptation.

#### *What Should Be Summarized?*

The increasing trend in both single-encounter and multiencounter summarizations suggests that both types are salient and should be used depending on the care situation.

Regarding the input for summarization, several experiments show that HCPs can act at least as accurately and in a timely manner with summarized structured [104–110,156] data or textual data [60,157] or with most information coded in these types of data [158] than using complete documentation. Therefore, the focus should be on summarizing textual and structured data when creating summaries for HCPs.

The increasing trend of using textual data for summarization might be attributed to the improvement of NLP, the improved computing power required for some NLP tasks, and the results published by Van Vleck et al [158], who claimed that a significant portion of patient information lies in clinical notes. In contrast, Hsu et al [111] challenged this hypothesis by presenting experiments to predict some clinical measures (eg, hospital readmission and mortality) using textual and structured patient information sources. They concluded that textual sources have little predictive power for the outcomes. However, their analysis might be biased by their methodology, as they use only simple syntactic metrics to describe textual information, whereas semantic information is not included in their model.

#### *How Data Can Be Summarized?*

The records analyzed in this review show myriad techniques for summarizing clinical data. Some are intrinsic to the input data type and work only with a particular data type, whereas others are not dependent on the input data type.

For textual data, the review reveals more works about abstractive summarization than extractive summarization or topic modeling combined, whereas in the general domain, topic modeling and

extractive summarization techniques are the most researched [58,159]. This discrepancy suggests that despite abstractive summarization techniques being immature [160], general problems with extractive summarization, such as redundancy [161], lack of coherence [162,163], and lengthiness [163], can be problematic for clinical applications. The verdict about topic modeling is unclear. Arnold et al [112] argue that clinicians are good at interpreting topic model results, but other records using this technique do not present evaluations with human participants.

An alternative (and natural) way of organizing summarization methods is to assess how they contribute to the summarization process. Motivated by the lack of a widely accepted theory of the summarization process, this review proposes a 3-step (collect—synthesize—communicate) framework to describe the summarization process, where each step should ideally be addressed by all summarization methods.

For the information collection step, many studies assume an easily queriable information source or propose medical concept extraction from textual data as a solution. More complex information (context, syntactic, or semantic structure of statements) is extracted in only a few studies, and some works propose extracting specific aspects as information.

Concerning information synthesis, a common approach is to precisely define the content of the summary (eg, [118,165,166]) or at least its format [167]. However, these studies do not evaluate the quality of their proposition (except the study by Ham et al [119]). In contrast, some records carried out experiments on the information needs of physicians [158,168]; however, the results were not integrated into any of the reviewed systems.

Concerning summary visualizations, there is no clear opinion on whether textual or graphical summaries are preferable in the medical context. Although there is a slight dominance of graphical displays among the analyzed records, some works [169,170] argue that textual summaries lead to more accurate decisions. However, a general pattern of these works is that they compare a specific graphical display with a particular textual display, limiting generalizability. These contradictory results suggest that both formats are helpful for clinical summarization, if relevant features are present. Problem-oriented views presented in some records (eg, [120]) can include both types of display and might have other advantages, as they group all available information about patient-specific problems [171].

Concerning the view on summarization, both top-down and bottom-up approaches are justifiable in a clinical setting. However, several bottom-up approaches have been inspired by studies that use top-down approaches. One example is the recent development of techniques for identifying salient concepts in source documents for abstractive summarization. This phenomenon may be due to the natural need for accountability and interpretability, which can be achieved more easily with a bottom-up approach closer to human cognition. The need for bottom-up approaches also suggests that there is a need that summarization techniques address all 3 steps of the proposed “collect—synthesize—communicate” framework, including information collection, synthesis, and visualization.

### ***How Are the Temporal, Uncertain Aspects of Information and Its Medical Persistence Conserved and Used?***

The temporal nature of clinical data is an essential aspect of clinical reasoning [172], and a relatively large portion of analyzed records presents solutions to use this aspect of information. However, in most of these studies, this aspect was only represented as a visualization feature. Most visualizations are timeline visualizations, plotting information along a horizontal or vertical temporal axis following Plaisant et al [175], although some alternative methods exist [121-124]. Alternatively, some studies have revealed the relationship between events by analyzing how variables change during the patient journey.

Although temporality in clinical settings is believed to be more complex [172] than a series of punctual events, current solutions to clinical summarization hardly reflect this complexity. Very few studies have attempted to incorporate more temporal information by using more complex models of temporality (eg, events lasting during an interval). Complex temporal information is usually not directly available in patient records and must be deduced from the context and knowledge-based rules. This process is called temporal abstraction and was applied in previous studies [90,91,129-132,176]. Hunter et al [129,130] considered the uncertainty of temporal information by defining the beginning and end of each time interval as an interval.

Although several levels of uncertainty exist in clinical care [177], the majority of the analyzed reports do not present solutions to conserve or handle any information uncertainty.

To a lesser extent, the pertinence of medical knowledge has been overlooked by many summarization approaches. Many records do not consider medical pertinence or use it only for some design considerations. However, the few records that handle this aspect of medical data provide a relatively wide range of solutions to constrain the resulting summaries. In most cases, these constraints are relatively weak; for example, concepts are assumed to be part of a specific medical ontology. This is obviously the case for concept extraction tools, but the records using reinforcement learning approximate factual correctness using this approach as well.

Deeper integration of medical knowledge is only present in works using medical rules to select salient information and in the 2 works using medical rules to create summaries. Liang et al [9] used medical knowledge to create components of their proposed NLP pipeline, whereas Shi et al [102] used medical knowledge-based rules to visualize abnormalities in the human body.

### ***How to Identify a Good Summary?***

Using the definition of clinical summarization (ie, simplifying and presenting information so that HCPs can act more smoothly and efficiently in different clinical situations), the ultimate purpose of an evaluation might be to determine whether using the proposed summarization systems would improve the efficiency of HCPs.

However, such an evaluation is often unfeasible owing to the high costs and ethical issues associated with potential medical errors.

This is supported by the results, as many of the proposed evaluations are approximative solutions, and there is quasi-uniform agreement that evaluating summarization is challenging and suboptimal [168]. The spectrum of these evaluations is broad, but a common trend is to carry out a qualitative evaluation using an easily calculable evaluation metric describing either the quality of the summary or its “usefulness” to perform a proxy task (ie, disease prediction).

These qualitative analyses are suboptimal. For example, one of the most common qualitative metrics, the ROUGE score, assumes a human-annotated “gold standard” summary to which to compare, but this standard may not exist given the high cost of annotation or because there are disagreements between people about what would be a “gold standard summary” [70,168]. To tackle this problem, some records [72,178] present a comparison between the semantic distribution of the input and the summary, whereas others [133,150,179] use heuristics to evaluate the results. Another problem with the ROUGE score is that even with a high ROUGE score, a summary can be very inaccurate [168]; therefore, there have been attempts to measure the “faithfulness” of summaries either by the number of medical concepts retrieved [81] or with a more complex faithfulness measure defined by Zhang et al [83].

Evaluations with human participants often complement the qualitative evaluations. Human evaluations have mainly positive outcomes (except in the study by van Amsterdam et al [148]); however, most of the evaluations are carried out on a small scale. This can explain why very few long-lasting implementations in health care have been presented in the literature.

It is also important to mention that there is very little comparison between summarization methods, and citations between records are scarce. This suggests that the research in this domain is fragmented.

These shortcomings suggest that evaluation is a weak point in clinical summarization proposals, and the lack of widely accepted evaluation metrics and methodology might be a main obstacle for research in the field.

### **Limitations**

Methodological biases are present in selection, synthesis, and reporting. First, the number of reviewers was limited both in the selection and analysis of records, resulting in selection and synthesis bias.

Selection bias also comes from the fact that the review was carried out on works published in a scientific paper or in the gray literature, and the initial search was carried out on 2 databases that are more specific to medicine. However, several unpublished summarization solutions have been applied to current EHR systems.

Moreover, publishing bias also adds to selection bias, as there is a clear dominance of positive results in scientific publishing.



Furthermore, the applied research queries and the selection of the database are also a source of bias. Using other data sources (eg, IEEE Xplore and Scopus) might have introduced further bias to the analysis. However, including the citations and references in the review process might have reduced this bias significantly. Moreover, the queries are formulated in English; therefore, results not in English and containing non-English terms might be missed if their abstract was not translated or if they do not appear in the citation list or references in the retrieved articles. Finally, there were potentially relevant records [149,180,181], where the full text could not be read and analyzed as it was not available at the time of writing the manuscript.

## Conclusions

Clinical summarization has not reached all domains at the same pace, although it is potentially beneficial in most medical fields. Two aspects, the requirement for quick decision-making and the overabundance of data, were identified as the main drivers for the development of automatic summarization methods. However, other less-evident drivers might also play a role in adaptation.

Despite this need, very few [113,119,182] scientific publications are presenting adaptation in real clinical settings, suggesting a low success rate in clinical environments.

Despite the large number and variety of propositions, hardly any comparisons exist between the solutions. This low rate is due to the difficulty in comparing the summarization methods.

From a cognitive psychological perspective and to measure how the summarization methods align with the definition of summarization, this review proposes to compare these algorithms through a “collect—synthesis—communicate” framework referring to information gathering from data, its synthesis, and communication to the end user.

Only a small proportion of the current propositions address all 3 steps, and none of the most abundant methods (ie, abstractive, extractive summarization, and visual design) address them completely.

Beyond the lack of alignment of the dimensions of summarization, propositions conserve and use crucial aspects of information (temporality, uncertainty, and medical pertinence) to varying extents.

Although uncertainty is rarely considered, temporality and some medical pertinence are conserved during some presentations, but the solutions are often preliminary or lack depth in these aspects. Further research is necessary to address these issues.

Nevertheless, the main shortcoming of the current automatic summarization methods is the lack of consistent evaluation. Although there are some new proposals to evaluate the quality of summarizations more rigorously [83], further research is required to relate these metrics to human perceptions.

## Acknowledgments

This work has been supported by “The Fondation privée des HUG” and the University Hospitals of Geneva. The authors would like to thank the reviewers for their valuable comments, which helped them improve the quality of the manuscript.

## Conflicts of Interest

CL is the editor-in-chief of the JMI Journal, but was not involved in the process of reviewing/accepting this paper. The author have no further interests to declare.

## Multimedia Appendix 1

PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist for the review.

[\[PDF File \(Adobe PDF File\), 497 KB-Multimedia Appendix 1\]](#)

## Multimedia Appendix 2

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram describing the review process applied in the study [182].

[\[DOCX File , 59 KB-Multimedia Appendix 2\]](#)

## Multimedia Appendix 3

Histogram showing the distribution of the number of evaluators in studies where evaluations with human participants are present.

[\[DOCX File , 53 KB-Multimedia Appendix 3\]](#)

## Multimedia Appendix 4

Histogram showing the number of other publications appearing in the review cited by each report. The number of cross-citations between the works are low.

[\[DOCX File , 50 KB-Multimedia Appendix 4\]](#)

## References

1. Alterman R. Understanding and summarization. *Artif Intell Rev* 1991;5(4):239-254 [[FREE Full text](#)] [doi: [10.1007/bf00141756](#)]
2. Feblowitz JC, Wright A, Singh H, Samal L, Sittig DF. Summarization of clinical information: a conceptual model. *J Biomed Inform* 2011 Aug;44(4):688-699 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2011.03.008](#)] [Medline: [21440086](#)]
3. Powsner SM, Tufte ER. Graphical summary of patient status. *Lancet* 1994 Aug 06;344(8919):386-389 [doi: [10.1016/s0140-6736\(94\)91406-0](#)] [Medline: [7914312](#)]
4. Hall A, Walton G. Information overload within the health care system: a literature review. *Health Info Libr J* 2004 Jun;21(2):102-108 [[FREE Full text](#)] [doi: [10.1111/j.1471-1842.2004.00506.x](#)] [Medline: [15191601](#)]
5. Botsis T, Hartvigsen G, Chen F, Weng C. Secondary use of EHR: data quality issues and informatics opportunities. *Summit Transl Bioinform* 2010 Mar 01;2010:1-5 [[FREE Full text](#)] [Medline: [21347133](#)]
6. Nijor S, Rallis G, Lad N, Gokcen E. Patient safety issues from information overload in electronic medical records. *J Patient Saf* 2022 Sep 01;18(6):e999-1003 [[FREE Full text](#)] [doi: [10.1097/PTS.0000000000001002](#)] [Medline: [35985047](#)]
7. Mamykina L, Vawdrey DK, Stetson PD, Zheng K, Hripcsak G. Clinical documentation: composition or synthesis? *J Am Med Inform Assoc* 2012 Nov;19(6):1025-1031 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2012-000901](#)] [Medline: [22813762](#)]
8. O'Donnell HC, Kaushal R, Barrón Y, Callahan MA, Adelman RD, Siegler EL. Physicians' attitudes towards copy and pasting in electronic note writing. *J Gen Intern Med* 2009 Jan 8;24(1):63-68 [[FREE Full text](#)] [doi: [10.1007/s11606-008-0843-2](#)] [Medline: [18998191](#)]
9. Liang JJ, Tsou CH, Dandala B, Poddar A, Joopudi V, Mahajan D, et al. Reducing physicians' cognitive load during chart review: a problem-oriented summary of the patient electronic record. *AMIA Annu Symp Proc* 2022 Feb 21;2021:763-772 [[FREE Full text](#)] [Medline: [35308927](#)]
10. Pivovarov R, Elhadad N. Automated methods for the summarization of electronic health records. *J Am Med Inform Assoc* 2015 Sep;22(5):938-947 [[FREE Full text](#)] [doi: [10.1093/jamia/ocv032](#)] [Medline: [25882031](#)]
11. Mishra R, Bian J, Fiszman M, Weir C, Jonnalagadda S, Mostafa J, et al. Text summarization in the biomedical domain: a systematic review of recent research. *J Biomed Inform* 2014 Dec;52:457-467 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2014.06.009](#)] [Medline: [25016293](#)]
12. Wang M, Wang M, Yu F, Yang Y, Walker J, Mostafa J. A systematic review of automatic text summarization for biomedical literature and EHRs. *J Am Med Inform Assoc* 2021 Sep 18;28(10):2287-2297 [[FREE Full text](#)] [doi: [10.1093/jamia/ocab143](#)] [Medline: [34338801](#)]
13. Afantenos S, Karkaletsis V, Stamatopoulos P. Summarization from medical documents: a survey. *Artif Intell Med* 2005 Feb;33(2):157-177 [doi: [10.1016/j.artmed.2004.07.017](#)] [Medline: [15811783](#)]
14. Li I, Pan J, Goldwasser J, Verma N, Wong WP, Nuzumlali MY, et al. Neural natural language processing for unstructured data in electronic health records: a review. *Comput Sci Rev* 2022 Nov;46:100511 [[FREE Full text](#)] [doi: [10.1016/j.cosrev.2022.100511](#)]
15. Koleček TA, Dreisbach C, Bourne PE, Bakken S. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. *J Am Med Inform Assoc* 2019 Apr 01;26(4):364-379 [[FREE Full text](#)] [doi: [10.1093/jamia/ocy173](#)] [Medline: [30726935](#)]
16. van Buchem MM, Boosman H, Bauer MP, Kant IM, Cammel SA, Steyerberg EW. The digital scribe in clinical practice: a scoping review and research agenda. *NPJ Digit Med* 2021 Mar 26;4(1):57 [[FREE Full text](#)] [doi: [10.1038/s41746-021-00432-5](#)] [Medline: [33772070](#)]
17. Boyd AD, Young CD, Amatayakul M, Dieter MG, Pawola LM. Developing visual thinking in the electronic health record. *Stud Health Technol Inform* 2017;245:308-312 [Medline: [29295105](#)]
18. Rind A, Federico P, Gschwandtner T, Aigner W, Doppler J, Wagner M. Visual analytics of electronic health records with a focus on time. In: Rinaldi G, editor. *New Perspectives in Medical Records: Meeting the Needs of Patients and Practitioners*. Cham, Switzerland: Springer; 2017:65-77
19. West VL, Borland D, Hammond WE. Innovative information visualization of electronic health record data: a systematic review. *J Am Med Inform Assoc* 2015 Mar;22(2):330-339 [[FREE Full text](#)] [doi: [10.1136/amiajnl-2014-002955](#)] [Medline: [25336597](#)]
20. Dowding D, Randell R, Gardner P, Fitzpatrick G, Dykes P, Favela J, et al. Dashboards for improving patient care: review of the literature. *Int J Med Inform* 2015 Feb;84(2):87-100 [[FREE Full text](#)] [doi: [10.1016/j.ijmedinf.2014.10.001](#)] [Medline: [25453274](#)]
21. Wright MC, Borbolla D, Waller RG, Del Fiol G, Reese T, Nesbitt P, et al. Critical care information display approaches and design frameworks: a systematic review and meta-analysis. *J Biomed Inform X* 2019 Sep;3:100041 [[FREE Full text](#)] [doi: [10.1016/j.yjbix.2019.100041](#)] [Medline: [31423485](#)]
22. Segall N, Borbolla D, Del FG, Waller R, Reese T, Nesbitt P, et al. Trend displays to support critical care: a systematic review. In: *Proceedings of the 2017 IEEE International Conference on Healthcare Informatics*. 2017 Presented at: ICHI '17; August 23-26, 2017; Park City, UT p. 305-313 URL: <https://ieeexplore.ieee.org/abstract/document/8031160>

23. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018 Oct 02;169(7):467-473 [[FREE Full text](#)] [doi: [10.7326/M18-0850](#)] [Medline: [30178033](#)]
24. Kellermeyer L, Harnke B, Knight S. Covidence and rayyan. *J Med Libr Assoc* 2018 Oct 04;106(4):580-583 [[FREE Full text](#)] [doi: [10.5195/jmla.2018.513](#)]
25. Nelson KE. What do you do if you can't tell the whole story? The development of summarization skills. In: Nelson KE, editor. *Children's Language*. Volume 4. New York, NY. Psychology Press; 1983:315-383
26. Hidi S, Anderson V. Producing written summaries: task demands, cognitive operations, and implications for instruction. *Rev Educ Res* 1986;56(4):473-493 [[FREE Full text](#)] [doi: [10.3102/00346543056004473](#)]
27. Spärck Jones K. Automatic summarising: the state of the art. *Inf Process Manag* 2007 Nov;43(6):1449-1481 [[FREE Full text](#)] [doi: [10.1016/j.ipm.2007.03.009](#)]
28. Jones K. Automatic summarizing: factors and directions. In: Mani I, Maybury MT, editors. *Advances in Automatic Text Summarization*. Cambridge, MA. MIT Press; 1999:1-12
29. Liang J, Tsou C, Poddar A. A novel system for extractive clinical note summarization using EHR data. In: *Proceedings of the 2nd Clinical Natural Language Processing Workshop*. 2019 Presented at: ClinicalNLP '19; June 5-7, 2019; Minneapolis, MN p. 46-54 URL: <https://aclanthology.org/W19-1906.pdf> [doi: [10.18653/v1/w19-1906](#)]
30. Gundogdu B, Pamuksuz U, Chung JH, Telleria JM, Liu P, Khan F, et al. Customized impression prediction from radiology reports using BERT and LSTMs. *IEEE Trans Artif Intell* 2023 Aug;4(4):744-753 [doi: [10.1109/tai.2021.3086435](#)]
31. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71 [[FREE Full text](#)] [doi: [10.1136/bmj.n71](#)] [Medline: [33782057](#)]
32. Abacha AB, Mrabet Y, Zhang Y, Shivade C, Langlotz C, Demner-Fushman D. Overview of the MEDIQA 2021 shared task on summarization in the medical domain. In: *Proceedings of the 20th Workshop on Biomedical Language Processing*. 2021 Presented at: BioNLP '21; June 11, 2021; Virtual Event p. 74-85 URL: <https://aclanthology.org/2021.bionlp-1.8.pdf> [doi: [10.18653/v1/2021.bionlp-1.8](#)]
33. Moen H, Peltonen LM, Heimonen J, Airola A, Pahikkala T, Salakoski T, et al. Comparison of automatic summarisation methods for clinical free text notes. *Artif Intell Med* 2016 Feb;67:25-37 [[FREE Full text](#)] [doi: [10.1016/j.artmed.2016.01.003](#)] [Medline: [26900011](#)]
34. Reunamo A, Peltonen P, Mustonen M, Saari M, Salakoski T, Salanterä S, et al. Text classification model explainability for keyword extraction - towards keyword-based summarization of nursing care episodes. *Stud Health Technol Inform* 2022 Jun 06;290:632-636 [[FREE Full text](#)] [doi: [10.3233/SHTI220154](#)] [Medline: [35673093](#)]
35. Deng Y, Denecke K. Visualizing unstructured patient data for assessing diagnostic and therapeutic history. *Stud Health Technol Inform* 2014;205:1158-1162 [Medline: [25160371](#)]
36. Liang S, Kades K, Fink M, Full P, Weber T, Kleesiek J, et al. Fine-tuning BERT models for summarizing German radiology findings. In: *Proceedings of the 4th Clinical Natural Language Processing Workshop*. 2022 Presented at: ClinicalNLP '22; July 14, 2022; Virtual Event p. 30-40 URL: <https://aclanthology.org/2022.clinicalnlp-1.4.pdf> [doi: [10.18653/v1/2022.clinicalnlp-1.4](#)]
37. Loukas C. Video content analysis of surgical procedures. *Surg Endosc* 2018 Feb 26;32(2):553-568 [doi: [10.1007/s00464-017-5878-1](#)] [Medline: [29075965](#)]
38. Koopman RJ, Mainous AG3. Evaluating multivariate risk scores for clinical decision making. *Fam Med* 2008 Jun;40(6):412-416 [Medline: [18773779](#)]
39. Miller RL, Vergara VM, Pearson GD, Calhoun VD. Multiframe Evolving Dynamic Functional Connectivity (EVOdFNC): a method for constructing and investigating functional brain motifs. *Front Neurosci* 2022 Apr 19;16:770468 [[FREE Full text](#)] [doi: [10.3389/fnins.2022.770468](#)] [Medline: [35516809](#)]
40. Hamza R, Muhammad K, Lv Z, Titouna F. Secure video summarization framework for personalized wireless capsule endoscopy. *Pervasive Mob Comput* 2017 Oct;41:436-450 [[FREE Full text](#)] [doi: [10.1016/j.pmcj.2017.03.011](#)]
41. Muhammad K, Ahmad J, Sajjad M, Baik SW. Visual saliency models for summarization of diagnostic hysteroscopy videos in healthcare systems. *Springerplus* 2016 Sep 6;5(1):1495 [[FREE Full text](#)] [doi: [10.1186/s40064-016-3171-8](#)] [Medline: [27652068](#)]
42. Ma M, Mei S, Wan S, Wang Z, Ge Z, Lam V, et al. Keyframe extraction from laparoscopic videos via diverse and weighted dictionary selection. *IEEE J Biomed Health Inform* 2021 May;25(5):1686-1698 [doi: [10.1109/JBHI.2020.3019198](#)] [Medline: [32841131](#)]
43. Yaesoubi M, Silva RF, Iraj A, Calhoun VD. Frequency-aware summarization of resting-state fMRI data. *Front Syst Neurosci* 2020 Apr 7;14:16 [[FREE Full text](#)] [doi: [10.3389/fnsys.2020.00016](#)] [Medline: [32317942](#)]
44. Huang R, Ying Q, Lin Z, Zheng Z, Tan L, Tang G, et al. Extracting keyframes of breast ultrasound video using deep reinforcement learning. *Med Image Anal* 2022 Aug;80:102490 [doi: [10.1016/j.media.2022.102490](#)] [Medline: [35717873](#)]
45. Hendryx EP, Rivière BM, Sorensen DC, Rusin CG. Finding representative electrocardiogram beat morphologies with CUR. *J Biomed Inform* 2018 Jan;77:97-110 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2017.12.003](#)] [Medline: [29224855](#)]

46. Park J, Kang K. HeartSearcher: finds patients with similar arrhythmias based on heartbeat classification. *IET Syst Biol* 2015 Dec;9(6):303-308 [[FREE Full text](#)] [doi: [10.1049/iet-syb.2015.0011](https://doi.org/10.1049/iet-syb.2015.0011)] [Medline: [26577165](#)]
47. Akalin A, Franke V, Vlahoviček K, Mason CE, Schübeler D. Genomation: a toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics* 2015 Apr 01;31(7):1127-1129 [doi: [10.1093/bioinformatics/btu775](https://doi.org/10.1093/bioinformatics/btu775)] [Medline: [25417204](#)]
48. Cai X, Liu S, Han J, Yang L, Liu Z, Liu T. ChestXRyBERT: a pretrained language model for chest radiology report summarization. *IEEE Trans Multimedia* 2021;25:845-855 [[FREE Full text](#)] [doi: [10.1109/tmm.2021.3132724](https://doi.org/10.1109/tmm.2021.3132724)]
49. MacAvaney S, Sotudeh S, Cohan A, Goharian N, Talati I, Filice R. Ontology-aware clinical abstractive summarization. In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2019 Presented at: SIGIR'19; July 21-25, 2019; Paris, France p. 1013-1016 URL: <https://dl.acm.org/doi/10.1145/3331184.3331319> [doi: [10.1145/3331184.3331319](https://doi.org/10.1145/3331184.3331319)]
50. Sotudeh S, Goharian N, Filice R. Attend to medical ontologies: content selection for clinical abstractive summarization. *arXiv. Preprint posted online May 1, 2020* 2020 [[FREE Full text](#)] [doi: [10.18653/v1/2020.acl-main.172](https://doi.org/10.18653/v1/2020.acl-main.172)]
51. Zhang Y, Ding D, Qian T, Manning C, Langlotz C. Learning to summarize radiology findings. *arXiv. Preprint posted online September 12, 2018* 2018 [[FREE Full text](#)] [doi: [10.18653/v1/w18-5623](https://doi.org/10.18653/v1/w18-5623)]
52. Manas G, Aribandi V, Kursuncu U, Alambo A, Shalin VL, Thirunarayan K, et al. Knowledge-infused abstractive summarization of clinical diagnostic interviews: framework development study. *JMIR Ment Health* 2021 May 10;8(5):e20865 [[FREE Full text](#)] [doi: [10.2196/20865](https://doi.org/10.2196/20865)] [Medline: [33970116](#)]
53. Chen YP, Chen YY, Lin JJ, Huang CH, Lai F. Modified bidirectional encoder representations from transformers extractive summarization model for hospital information systems based on character-level tokens (AlphaBERT): development and performance evaluation. *JMIR Med Inform* 2020 Apr 29;8(4):e17787 [[FREE Full text](#)] [doi: [10.2196/17787](https://doi.org/10.2196/17787)] [Medline: [32347806](#)]
54. Gaut G, Steyvers M, Imel ZE, Atkins DC, Smyth P. Content coding of psychotherapy transcripts using labeled topic models. *IEEE J Biomed Health Inform* 2017 Mar;21(2):476-487 [[FREE Full text](#)] [doi: [10.1109/JBHI.2015.2503985](https://doi.org/10.1109/JBHI.2015.2503985)] [Medline: [26625437](#)]
55. Speier W, Ong MK, Arnold CW. Using phrases and document metadata to improve topic modeling of clinical reports. *J Biomed Inform* 2016 Jun;61:260-266 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2016.04.005](https://doi.org/10.1016/j.jbi.2016.04.005)] [Medline: [27109931](#)]
56. Shah J, Mohammed S. Clinical narrative summarization based on the MIMIC III dataset. *J Multimedia Ubiquitous Eng* 2020 Nov;15(2):60 [[FREE Full text](#)] [doi: [10.21742/IJMUE.2020.15.2.05](https://doi.org/10.21742/IJMUE.2020.15.2.05)]
57. Chen JH, Goldstein MK, Asch SM, Mackey L, Altman RB. Predicting inpatient clinical order patterns with probabilistic topic models vs conventional order sets. *J Am Med Inform Assoc* 2017 May 01;24(3):472-480 [[FREE Full text](#)] [doi: [10.1093/jamia/ocw136](https://doi.org/10.1093/jamia/ocw136)] [Medline: [27655861](#)]
58. Allahyari M, Pouriyeh S, Assefi M, Safaei S, D. ED, B. JB, et al. Text summarization techniques: a brief survey. *Int J Adv Comput Sci Appl* 2017;8(10):397-405 [[FREE Full text](#)] [doi: [10.14569/ijacsa.2017.081052](https://doi.org/10.14569/ijacsa.2017.081052)]
59. Jane Y, Nehemiah HK, Kannan A. Classifying unevenly spaced clinical time series data using forecast error approximation based bottom-up (FeAB) segmented time delay neural network. *Comput Methods Biomech Biomed Eng Imaging Vis* 2020 Dec 21;9(1):92-105 [doi: [10.1080/21681163.2020.1817791](https://doi.org/10.1080/21681163.2020.1817791)]
60. Portet F, Reiter E, Gatt A, Hunter J, Sripada S, Freer Y, et al. Automatic generation of textual summaries from neonatal intensive care data. *Artif Intell* 2009 May;173(7-8):789-816 [[FREE Full text](#)] [doi: [10.1016/j.artint.2008.12.002](https://doi.org/10.1016/j.artint.2008.12.002)]
61. Hsu W, Taira RK. Tools for improving the characterization and visualization of changes in neuro-oncology patients. *AMIA Annu Symp Proc* 2010 Nov 13;2010:316-320 [[FREE Full text](#)] [Medline: [21346992](#)]
62. Albers D, Elhadad N, Claassen J, Perotte R, Goldstein A, Hripesak G. Estimating summary statistics for electronic health record laboratory data for use in high-throughput phenotyping algorithms. *J Biomed Inform* 2018 Feb;78:87-101 [[FREE Full text](#)] [doi: [10.1016/j.jbi.2018.01.004](https://doi.org/10.1016/j.jbi.2018.01.004)] [Medline: [29369797](#)]
63. Landi I, Glicksberg BS, Lee H, Cherng S, Landi G, Danieleto M, et al. Deep representation learning of electronic health records to unlock patient stratification at scale. *NPJ Digit Med* 2020 Jul 17;3(1):96 [[FREE Full text](#)] [doi: [10.1038/s41746-020-0301-z](https://doi.org/10.1038/s41746-020-0301-z)] [Medline: [32699826](#)]
64. Miotto R, Li L, Kidd B, Dudley J. Deep patient: an unsupervised representation to predict the future of patients from the electronic health records. *Sci Rep* 2016 May 17;6:26094 [[FREE Full text](#)] [doi: [10.1038/srep26094](https://doi.org/10.1038/srep26094)] [Medline: [27185194](#)]
65. Lewis M, Liu Y, Goyal N, Ghazvininejad M, Mohamed A, Levy O, et al. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv. Preprint posted online October 29, 2019* 2019 [[FREE Full text](#)] [doi: [10.18653/v1/2020.acl-main.703](https://doi.org/10.18653/v1/2020.acl-main.703)]
66. Devlin J, Chang M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. *arXiv. Preprint posted online October 11, 2018* 2018 [[FREE Full text](#)]
67. Zhang J, Zhao Y, Saleh M, Liu P. PEGASUS: pre-training with extracted gap-sentences for abstractive summarization. In: *Proceedings of the 37th International Conference on Machine Learning*. 2020 Presented at: ICML '20; July 13-18, 2020; Vienna, Austria p. 11328-11339 URL: <https://dl.acm.org/doi/abs/10.5555/3524938.3525989> [doi: [10.5555/3524938.3525989](https://doi.org/10.5555/3524938.3525989)]
68. Ribeiro M, Singh S, Guestrin C. "Why should I trust you?": Explaining the predictions of any classifier. *arXiv. Preprint posted online February 16, 2016* 2016 Aug 9 [[FREE Full text](#)] [doi: [10.1145/2939672.2939778](https://doi.org/10.1145/2939672.2939778)]



69. Subramanian V, Engelhard M, Berchuck S, Chen L, Henao R, Carin L. SpanPredict: extraction of predictive document spans with neural attention. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language. 2021 Presented at: NAACL '21; June 6-11, 2021; Stroudsburg, PA p. 5234-5258 URL: <https://aclanthology.org/2021.naacl-main.413.pdf> [doi: [10.18653/v1/2021.naacl-main.413](https://doi.org/10.18653/v1/2021.naacl-main.413)]
70. Alsentzer E, Kim A. Extractive summarization of ehr discharge notes. arXiv. Preprint posted online October 26, 2018 2018 [FREE Full text]
71. Liu X, Xu K, Xie P, Xing E. Unsupervised pseudo-labeling for extractive summarization on electronic health records. arXiv. Preprint posted online November 20, 2018 2018 [FREE Full text]
72. Kanwal N, Rizzo G. Attention-based clinical note summarization. In: Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing. 2022 Presented at: SAC '22; April 25-29, 2022; Virtual Event p. 813-820 URL: <https://dl.acm.org/doi/abs/10.1145/3477314.3507256> [doi: [10.1145/3477314.3507256](https://doi.org/10.1145/3477314.3507256)]
73. McInerney DJ, Dabiri B, Touret AS, Young G, Meent JW, Wallace B. Query-focused EHR summarization to aid imaging diagnosis. Proc Mach Learn Res 2020:632-659 [FREE Full text]
74. Hu J, Li J, Chen Z, Shen Y, Song Y, Wan X. Word graph guided summarization for radiology findings. arXiv. Preprint posted online December 18, 2021 2021 [FREE Full text] [doi: [10.18653/v1/2021.findings-acl.441](https://doi.org/10.18653/v1/2021.findings-acl.441)]
75. Hu J, Li Z, Chen Z, Li Z, Wan X. Graph enhanced contrastive learning for radiology findings summarization. arXiv. Preprint posted online April 1, 2022 2022 [FREE Full text] [doi: [10.18653/v1/2022.acl-long.320](https://doi.org/10.18653/v1/2022.acl-long.320)]
76. Chang D, Lin E, Brandt C, Taylor RA. Incorporating domain knowledge into language models by using graph convolutional networks for assessing semantic textual similarity: model development and performance comparison. JMIR Med Inform 2021 Nov 26;9(11):e23101 [FREE Full text] [doi: [10.2196/23101](https://doi.org/10.2196/23101)] [Medline: [34842531](https://pubmed.ncbi.nlm.nih.gov/34842531/)]
77. Mahajan D, Tsou CH, Liang JL. IBMResearch at MEDIQA 2021: toward improving factual correctness of radiology report abstractive summarization. In: Proceedings of the 20th Workshop on Biomedical Language Processing. 2021 Presented at: BioNLP '21; June 11, 2021; Virtual Event p. 302-310 URL: <https://aclanthology.org/2021.bionlp-1.35.pdf> [doi: [10.18653/v1/2021.bionlp-1.35](https://doi.org/10.18653/v1/2021.bionlp-1.35)]
78. Dai S, Wang Q, Lyu Y, Zhu Y. BDKG at MEDIQA 2021: system report for the radiology report summarization task. In: Proceedings of the 20th Workshop on Biomedical Language Processing. 2021 Presented at: BioNLP '21; June 11, 2021; Virtual Event p. 103-111 URL: <https://aclanthology.org/2021.bionlp-1.11.pdf> [doi: [10.18653/v1/2021.bionlp-1.11](https://doi.org/10.18653/v1/2021.bionlp-1.11)]
79. He Y, Chen M, Huang S. damo\_nlp at MEDIQA 2021: knowledge-based preprocessing and coverage-oriented reranking for medical question summarization. In: Proceedings of the 20th Workshop on Biomedical Language Processing. 2021 Presented at: BioNLP '21; June 11, 2021; Virtual Event p. 2112-2118 URL: <https://aclanthology.org/2021.bionlp-1.12.pdf> [doi: [10.18653/v1/2021.bionlp-1.12](https://doi.org/10.18653/v1/2021.bionlp-1.12)]
80. Kondadadi R, Manch S, Ngo J, McCormack R. Optum at MEDIQA 2021: abstractive summarization of radiology reports using simple BART finetuning. In: Proceedings of the 20th Workshop on Biomedical Language Processing. 2021 Presented at: BioNLP '20; June 11, 2021; Virtual Event p. 280-284 URL: <https://aclanthology.org/2021.bionlp-1.32.pdf> [doi: [10.18653/v1/2021.bionlp-1.32](https://doi.org/10.18653/v1/2021.bionlp-1.32)]
81. Shing H, Shivade C, Pourdamghani N, Nan F, Resnik P, Oard D, et al. Towards clinical encounter summarization: learning to compose discharge summaries from prior notes. arXiv. Preprint posted online April 27, 2021 2021 [FREE Full text]
82. Xu L, Zhang Y, Hong L, Cai Y, Sung S. ChicHealth@ MEDIQA 2021xploring the limits of pre-trained seq2seq models for medical summarization. In: Proceedings of the 20th Workshop on Biomedical Language Processing. 2021 Presented at: BioNLP '21; June 11, 2021; Virtual Event p. 263-267 URL: <https://aclanthology.org/2021.bionlp-1.29.pdf> [doi: [10.18653/v1/2021.bionlp-1.29](https://doi.org/10.18653/v1/2021.bionlp-1.29)]
83. Zhang Y, Merck D, Tsai EB, Manning CD, Langlotz CP. Optimizing the factual correctness of a summary: a study of summarizing radiology reports. arXiv. Preprint posted online November 6, 2019 2019 [FREE Full text] [doi: [10.18653/v1/2020.acl-main.458](https://doi.org/10.18653/v1/2020.acl-main.458)]
84. Delbrouck J, Zhang C, Rubin D. QIAI at MEDIQA 2021: multimodal radiology report summarization. In: Proceedings of the 20th Workshop on Biomedical Language Processing. 2021 Presented at: BioNLP '21; June 11, 2021; Virtual Event p. 285-290 URL: <https://aclanthology.org/2021.bionlp-1.33.pdf> [doi: [10.18653/v1/2021.bionlp-1.33](https://doi.org/10.18653/v1/2021.bionlp-1.33)]
85. Xu B, Gil-Jardiné C, Thiessard F, Tellier E, Avalos-Fernandez M, Lagarde E. Pre-training a neural language model improves the sample efficiency of an emergency room classification model. arXiv. Preprint posted online August 30, 2019 2019 [FREE Full text]
86. Tang L, Kooragayalu S, Wang Y, Ding Y, Durrett G, Rousseau JF, et al. EchoGen: a new benchmark study on generating conclusions from echocardiogram notes. Proc Conf Assoc Comput Linguist Meet 2022 May;2022:359-368 [FREE Full text] [doi: [10.18653/v1/2022.bionlp-1.35](https://doi.org/10.18653/v1/2022.bionlp-1.35)] [Medline: [36339656](https://pubmed.ncbi.nlm.nih.gov/36339656/)]
87. Lee E, Uppal K. CERC: an interactive content extraction, recognition, and construction tool for clinical and biomedical text. BMC Med Inform Decis Mak 2020 Dec 15;20(Suppl 14):306 [FREE Full text] [doi: [10.1186/s12911-020-01330-8](https://doi.org/10.1186/s12911-020-01330-8)] [Medline: [33323109](https://pubmed.ncbi.nlm.nih.gov/33323109/)]
88. Devarakonda MV, Mehta N, Tsou C, Liang JJ, Nowacki AS, Jelovsek JE. Automated problem list generation and physicians perspective from a pilot study. Int J Med Inform 2017 Sep;105:121-129 [FREE Full text] [doi: [10.1016/j.ijmedinf.2017.05.015](https://doi.org/10.1016/j.ijmedinf.2017.05.015)] [Medline: [28750905](https://pubmed.ncbi.nlm.nih.gov/28750905/)]

89. Zhu W, He Y, Chai L, Fan Y, Ni Y, Xie G, et al. paht\_nlp @ MEDIQA 2021: multi-grained query focused multi-answer summarization. In: Proceedings of the 20th Workshop on Biomedical Language Processing. 2021 Presented at: BioNLP '21; June 11, 2021; Virtual Event p. 966 URL: <https://aclanthology.org/2021.bionlp-1.10.pdf> [doi: [10.18653/v1/2021.bionlp-1.10](https://doi.org/10.18653/v1/2021.bionlp-1.10)]
90. Goldstein A, Shahar Y. An automated knowledge-based textual summarization system for longitudinal, multivariate clinical data. *J Biomed Inform* 2016 Jun;61:159-175 [FREE Full text] [doi: [10.1016/j.jbi.2016.03.022](https://doi.org/10.1016/j.jbi.2016.03.022)] [Medline: [27039119](https://pubmed.ncbi.nlm.nih.gov/27039119/)]
91. Goldstein A, Shahar Y, Orenbuch E, Cohen MJ. Evaluation of an automated knowledge-based textual summarization system for longitudinal clinical data, in the intensive care domain. *Artif Intell Med* 2017 Oct;82:20-33 [doi: [10.1016/j.artmed.2017.09.001](https://doi.org/10.1016/j.artmed.2017.09.001)] [Medline: [28958803](https://pubmed.ncbi.nlm.nih.gov/28958803/)]
92. Van Vleck TT, Elhadad N. Corpus-based problem selection for EHR note summarization. *AMIA Annu Symp Proc* 2010 Nov 13;2010:817-821 [FREE Full text] [Medline: [21347092](https://pubmed.ncbi.nlm.nih.gov/21347092/)]
93. Müller H, Reihs R, Posch A, Kremer A, Ulrich D, Zatloukal K. Data driven GUI design and visualization for a NGS based clinical decision support system. In: Proceedings of the 2016 20th International Conference Information Visualisation. 2016 Presented at: IV '16; July 19-22, 2016; Lisbon, Portugal p. 355-360 URL: <https://ieeexplore.ieee.org/abstract/document/7557952> [doi: [10.1109/iv.2016.79](https://doi.org/10.1109/iv.2016.79)]
94. Goff DJ, Loehfelm TW. Automated radiology report summarization using an open-source natural language processing pipeline. *J Digit Imaging* 2018 Apr;31(2):185-192 [FREE Full text] [doi: [10.1007/s10278-017-0030-2](https://doi.org/10.1007/s10278-017-0030-2)] [Medline: [29086081](https://pubmed.ncbi.nlm.nih.gov/29086081/)]
95. Kim B, Merchant M, Zheng C, Thomas A, Contreras R, Jacobsen S, et al. A natural language processing program effectively extracts key pathologic findings from radical prostatectomy reports. *J Endourol* 2014 Dec;28(12):1474-1478 [doi: [10.1089/end.2014.0221](https://doi.org/10.1089/end.2014.0221)] [Medline: [25211697](https://pubmed.ncbi.nlm.nih.gov/25211697/)]
96. Moradi M, Ghadiri N. Different approaches for identifying important concepts in probabilistic biomedical text summarization. *Artif Intell Med* 2018 Jan;84:101-116 [doi: [10.1016/j.artmed.2017.11.004](https://doi.org/10.1016/j.artmed.2017.11.004)] [Medline: [29208328](https://pubmed.ncbi.nlm.nih.gov/29208328/)]
97. Mane K, Bizon C, Schmitt C, Owen P, Burchett B, Pietrobon R, et al. VisualDecisionLinc: a visual analytics approach for comparative effectiveness-based clinical decision support in psychiatry. *J Biomed Inform* 2012 Feb;45(1):101-106 [FREE Full text] [doi: [10.1016/j.jbi.2011.09.003](https://doi.org/10.1016/j.jbi.2011.09.003)] [Medline: [21963813](https://pubmed.ncbi.nlm.nih.gov/21963813/)]
98. Mathe JL, Martin JB, Miller P, Ledeczi A, Weavind LM, Nadas A, et al. A model-integrated, guideline-driven, clinical decision-support system. *IEEE Softw* 2009 Jul;26(4):54-61 [FREE Full text] [doi: [10.1109/ms.2009.84](https://doi.org/10.1109/ms.2009.84)]
99. Semler M, Weavind L, Hooper M, Rice T, Gowda S, Nadas A, et al. An electronic tool for the evaluation and treatment of sepsis in the ICU: a randomized controlled trial. *Crit Care Med* 2015 Aug;43(8):1595-1602 [FREE Full text] [doi: [10.1097/CCM.0000000000001020](https://doi.org/10.1097/CCM.0000000000001020)] [Medline: [25867906](https://pubmed.ncbi.nlm.nih.gov/25867906/)]
100. Tignanelli CJ, Silverman GM, Lindemann EA, Trembley AL, Gipson JC, Beilman G, et al. Natural language processing of prehospital emergency medical services trauma records allows for automated characterization of treatment appropriateness. *J Trauma Acute Care Surg* 2020 May;88(5):607-614 [FREE Full text] [doi: [10.1097/TA.0000000000002598](https://doi.org/10.1097/TA.0000000000002598)] [Medline: [31977990](https://pubmed.ncbi.nlm.nih.gov/31977990/)]
101. Klumpner TT, Kountanis JA, Langen ES, Smith RD, Tremper KK. Use of a novel electronic maternal surveillance system to generate automated alerts on the labor and delivery unit. *BMC Anesthesiol* 2018 Jun 26;18(1):78 [FREE Full text] [doi: [10.1186/s12871-018-0540-6](https://doi.org/10.1186/s12871-018-0540-6)] [Medline: [29945569](https://pubmed.ncbi.nlm.nih.gov/29945569/)]
102. Shi L, Sun J, Yang Y, Ling T, Wang M, Gu Y, et al. Three-dimensional visual patient based on electronic medical diagnostic records. *IEEE J Biomed Health Inform* 2018 Jan;22(1):161-172 [doi: [10.1109/JBHI.2017.2702201](https://doi.org/10.1109/JBHI.2017.2702201)] [Medline: [28500014](https://pubmed.ncbi.nlm.nih.gov/28500014/)]
103. Hsu W, Taira R, El-Saden S, Kangaroo H, Bui A. Context-based electronic health record: toward patient specific healthcare. *IEEE Trans Inf Technol Biomed* 2012 Mar;16(2):228-234 [FREE Full text] [doi: [10.1109/TITB.2012.2186149](https://doi.org/10.1109/TITB.2012.2186149)] [Medline: [22395637](https://pubmed.ncbi.nlm.nih.gov/22395637/)]
104. Ahmed A, Chandra S, Herasevich V, Gajic O, Pickering BW. The effect of two different electronic health record user interfaces on intensive care provider task load, errors of cognition, and performance. *Crit Care Med* 2011 Jul;39(7):1626-1634 [doi: [10.1097/CCM.0b013e31821858a0](https://doi.org/10.1097/CCM.0b013e31821858a0)] [Medline: [21478739](https://pubmed.ncbi.nlm.nih.gov/21478739/)]
105. Foraker R, Kite B, Kelley M, Lai A, Roth C, Lopetegui M, et al. EHR-based visualization tool: adoption rates, satisfaction, and patient outcomes. *EGEMS (Wash DC)* 2015;3(2):1159 [FREE Full text] [doi: [10.13063/2327-9214.1159](https://doi.org/10.13063/2327-9214.1159)] [Medline: [26290891](https://pubmed.ncbi.nlm.nih.gov/26290891/)]
106. Kheterpal S, Shanks A, Tremper K. Impact of a novel multiparameter decision support system on intraoperative processes of care and postoperative outcomes. *Anesthesiology* 2018 Feb;128(2):272-282 [FREE Full text] [doi: [10.1097/ALN.0000000000002023](https://doi.org/10.1097/ALN.0000000000002023)] [Medline: [29337743](https://pubmed.ncbi.nlm.nih.gov/29337743/)]
107. Thursky KA, Mahemoff M. User-centered design techniques for a computerised antibiotic decision support system in an intensive care unit. *Int J Med Inform* 2007 Oct;76(10):760-768 [doi: [10.1016/j.ijmedinf.2006.07.011](https://doi.org/10.1016/j.ijmedinf.2006.07.011)] [Medline: [16950650](https://pubmed.ncbi.nlm.nih.gov/16950650/)]
108. Nelson O, Sturgis B, Gilbert K, Henry E, Clegg K, Tan JM, et al. A visual analytics dashboard to summarize serial anesthesia records in pediatric radiation treatment. *Appl Clin Inform* 2019 Aug 07;10(4):563-569 [FREE Full text] [doi: [10.1055/s-0039-1693712](https://doi.org/10.1055/s-0039-1693712)] [Medline: [31390667](https://pubmed.ncbi.nlm.nih.gov/31390667/)]
109. Monico LB, Ludwig A, Lertch E, Mitchell SG. Using timeline methodology to visualize treatment trajectories of youth and young adults following inpatient opioid treatment. *Int J Qual Methods* 2020 Dec;19:160940692097010 [FREE Full text] [doi: [10.1177/1609406920970106](https://doi.org/10.1177/1609406920970106)]



110. Miller A, Scheinkestel C, Steele C. The effects of clinical information presentation on physicians' and nurses' decision-making in ICUs. *Appl Ergon* 2009 Jul;40(4):753-761 [doi: [10.1016/j.apergo.2008.07.004](https://doi.org/10.1016/j.apergo.2008.07.004)] [Medline: [18834970](#)]
111. Hsu CC, Karnwal S, Mullainathan S, Obermeyer Z, Tan C. Characterizing the value of information in medical notes. arXiv. Preprint posted online December 9, 2020 2020 [FREE Full text] [doi: [10.18653/v1/2020.findings-emnlp.187](https://doi.org/10.18653/v1/2020.findings-emnlp.187)]
112. Arnold CW, Oh A, Chen S, Speier W. Evaluating topic model interpretability from a primary care physician perspective. *Comput Methods Programs Biomed* 2016 Feb;124:67-75 [FREE Full text] [doi: [10.1016/j.cmpb.2015.10.014](https://doi.org/10.1016/j.cmpb.2015.10.014)] [Medline: [26614020](#)]
113. Hirsch J, Tanenbaum J, Lipsky Gorman S, Liu C, Schmitz E, Hashorva D, et al. HARVEST, a longitudinal patient record summarizer. *J Am Med Inform Assoc* 2015 Mar;22(2):263-274 [FREE Full text] [doi: [10.1136/amiajnl-2014-002945](https://doi.org/10.1136/amiajnl-2014-002945)] [Medline: [25352564](#)]
114. Dudko A, er; Endrjukaite T, Kiyoki Y. Medical documents processing for summary generation and keywords highlighting based on natural language processing and ontology graph descriptor approach. In: Proceedings of the 19th International Conference on Information Integration and Web-based Applications & Services. 2017 Presented at: iiWAS '17; December 4-6, 2017; Salzburg, Austria p. 58-65 URL: <https://dl.acm.org/doi/10.1145/3151759.3151784> [doi: [10.1145/3151759.3151784](https://doi.org/10.1145/3151759.3151784)]
115. Baldwin T, Guo Y, Mukherjee VV, Syeda-Mahmood T. Generalized extraction and classification of span-level clinical phrases. *AMIA Annu Symp Proc* 2018;2018:205-214 [FREE Full text] [Medline: [30815058](#)]
116. Kenei J, Opiyo E, Oboko R. Visualizing semantic structure of a clinical text document. *Eur J Electr Eng Comput Sci* 2020 Dec 04;4(6) [FREE Full text] [doi: [10.24018/ejece.2020.4.6.256](https://doi.org/10.24018/ejece.2020.4.6.256)]
117. Dagliati A, Sacchi L, Tibollo V, Cogni G, Teliti M, Martinez-Millana A, et al. A dashboard-based system for supporting diabetes care. *J Am Med Inform Assoc* 2018 May 01;25(5):538-547 [FREE Full text] [doi: [10.1093/jamia/ocx159](https://doi.org/10.1093/jamia/ocx159)] [Medline: [29409033](#)]
118. Carr LL, Zelarney P, Meadows S, Kern JA, Long MB, Kern E. Development of a cancer care summary through the electronic health record. *J Oncol Pract* 2016 Feb;12(2):e231-e240 [doi: [10.1200/jop.2015.006890](https://doi.org/10.1200/jop.2015.006890)]
119. Ham PB, Anderton T, Gallaher R, Hyrman M, Simmerman E, Ramanathan A, et al. Development of electronic medical record-based "rounds report" results in improved resident efficiency, more time for direct patient care and education, and less resident duty hour violations. *Am Surg* 2016 Sep 01;82(9):853-859 [FREE Full text] [doi: [10.1177/000313481608200950](https://doi.org/10.1177/000313481608200950)]
120. Klann J, McCoy A, Wright A, Wattanasin N, Sittig D, Murphy S. Health care transformation through collaboration on open-source informatics projects: integrating a medical applications platform, research data repository, and patient summarization. *Interact J Med Res* 2013 May 30;2(1):e11 [FREE Full text] [doi: [10.2196/ijmr.2454](https://doi.org/10.2196/ijmr.2454)] [Medline: [23722634](#)]
121. Flohr L, Beaudry S, Johnson KT, West N, Burns CM, Ansermino JM, et al. Clinician-driven design of VitalPAD-an intelligent monitoring and communication device to improve patient safety in the intensive care unit. *IEEE J Transl Eng Health Med* 2018 Mar 05;6:3000114 [FREE Full text] [doi: [10.1109/JTEHM.2018.2812162](https://doi.org/10.1109/JTEHM.2018.2812162)] [Medline: [29552425](#)]
122. Guo S, Xu K, Zhao R, Gotz D, Zha H, Cao N. EventThread: visual summarization and stage analysis of event sequence data. *IEEE Trans Vis Comput Graph* 2018 Jan;24(1):56-65 [doi: [10.1109/TVCG.2017.2745320](https://doi.org/10.1109/TVCG.2017.2745320)] [Medline: [28866586](#)]
123. Monroe M, Lan R, Lee H, Plaisant C, Shneiderman B. Temporal event sequence simplification. *IEEE Trans Vis Comput Graph* 2013 Dec;19(12):2227-2236 [doi: [10.1109/TVCG.2013.200](https://doi.org/10.1109/TVCG.2013.200)] [Medline: [24051789](#)]
124. Klimov D, Shahar Y, Taieb-Maimon M. Intelligent visualization and exploration of time-oriented data of multiple patients. *Artif Intell Med* 2010 May;49(1):11-31 [doi: [10.1016/j.artmed.2010.02.001](https://doi.org/10.1016/j.artmed.2010.02.001)] [Medline: [20303245](#)]
125. Brich N, Schulz C, Peter J, Klingert W, Schenk M, Weiskopf D, et al. Visual analytics of multivariate intensive care time series data. *Comput Graph Forum* 2022 Apr 13;41(6):273-286 [FREE Full text] [doi: [10.1111/cgf.14498](https://doi.org/10.1111/cgf.14498)]
126. Salvi E, Bosoni P, Tibollo V, Kruijver L, Calcaterra V, Sacchi L, et al. Patient-generated health data integration and advanced analytics for diabetes management: the AID-GM platform. *Sensors (Basel)* 2019 Dec 24;20(1):128 [FREE Full text] [doi: [10.3390/s20010128](https://doi.org/10.3390/s20010128)] [Medline: [31878195](#)]
127. Drews FA, Doig A. Evaluation of a configural vital signs display for intensive care unit nurses. *Hum Factors* 2014 May;56(3):569-580 [doi: [10.1177/0018720813499367](https://doi.org/10.1177/0018720813499367)] [Medline: [24930176](#)]
128. Sacchi L, Capozzi D, Bellazzi R, Larizza C. JTSA: an open source framework for time series abstractions. *Comput Methods Programs Biomed* 2015 Oct;121(3):175-188 [doi: [10.1016/j.cmpb.2015.05.006](https://doi.org/10.1016/j.cmpb.2015.05.006)] [Medline: [26120073](#)]
129. Hunter J, Freer Y, Gatt A, Logie R, McIntosh N, van der Meulen M, et al. Summarising complex ICU data in natural language. *AMIA Annu Symp Proc* 2008 Nov 06;2008:323-327 [FREE Full text] [Medline: [18998961](#)]
130. Hunter J, Freer Y, Gatt A, Reiter E, Sripada S, Sykes C. Automatic generation of natural language nursing shift summaries in neonatal intensive care: BT-Nurse. *Artif Intell Med* 2012 Nov;56(3):157-172 [doi: [10.1016/j.artmed.2012.09.002](https://doi.org/10.1016/j.artmed.2012.09.002)] [Medline: [23068882](#)]
131. Scott D, Hallett C, Fettiplace R. Data-to-text summarisation of patient records: using computer-generated summaries to access patient histories. *Patient Educ Couns* 2013 Aug;92(2):153-159 [FREE Full text] [doi: [10.1016/j.pec.2013.04.019](https://doi.org/10.1016/j.pec.2013.04.019)] [Medline: [23746770](#)]
132. Shahar Y, Goren-Bar D, Boaz D, Tahan G. Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions. *Artif Intell Med* 2006 Oct;38(2):115-135 [doi: [10.1016/j.artmed.2005.03.001](https://doi.org/10.1016/j.artmed.2005.03.001)] [Medline: [16343873](#)]

133. Levy-Fix G, Zucker J, Stojanovic K, Elhadad N. Towards patient record summarization through joint phenotype learning in HIV patients. arXiv. Preprint posted online March 9, 2020 2020 [FREE Full text]
134. Byrne CA, O'Grady M, Collier R, O'Hare GM. An evaluation of graphical formats for the summary of activities of daily living (ADLs). *Healthcare (Basel)* 2020 Jul 01;8(3):194 [FREE Full text] [doi: [10.3390/healthcare8030194](https://doi.org/10.3390/healthcare8030194)] [Medline: [32630304](https://pubmed.ncbi.nlm.nih.gov/32630304/)]
135. Zhang Y, Chanana K, Dunne C. IDMMVis: temporal event sequence visualization for type 1 diabetes treatment decision support. *IEEE Trans Vis Comput Graph (Forthcoming)* 2018 Aug 20 [doi: [10.1109/TVCG.2018.2865076](https://doi.org/10.1109/TVCG.2018.2865076)] [Medline: [30136981](https://pubmed.ncbi.nlm.nih.gov/30136981/)]
136. Welch G, Balder A, Zagarins S. Telehealth program for type 2 diabetes: usability, satisfaction, and clinical usefulness in an urban community health center. *Telemed J E Health* 2015 May;21(5):395-403 [doi: [10.1089/tmj.2014.0069](https://doi.org/10.1089/tmj.2014.0069)] [Medline: [25748544](https://pubmed.ncbi.nlm.nih.gov/25748544/)]
137. Sultanum N, Singh D, Brudno M, Chevalier F. Doccurate: a curation-based approach for clinical text visualization. *IEEE Trans Vis Comput Graph (Forthcoming)* 2018 Aug 20 [doi: [10.1109/TVCG.2018.2864905](https://doi.org/10.1109/TVCG.2018.2864905)] [Medline: [30136959](https://pubmed.ncbi.nlm.nih.gov/30136959/)]
138. Stubbs B, Kale D, Das A. Sim•TwentyFive: an interactive visualization system for data-driven decision support. *AMIA Annu Symp Proc* 2012;2012:891-900 [FREE Full text] [Medline: [23304364](https://pubmed.ncbi.nlm.nih.gov/23304364/)]
139. Lamy J, Duclos C, Hamek S, Beuscart-Zéphir MC, Kerdellhué G, Darmoni S, et al. Towards iconic language for patient records, drug monographs, guidelines and medical search engines. *Stud Health Technol Inform* 2010;160(Pt 1):156-160 [Medline: [20841669](https://pubmed.ncbi.nlm.nih.gov/20841669/)]
140. Albert R, Agutter J, Syroid N, Johnson K, Loeb R, Westenskow D. A simulation-based evaluation of a graphic cardiovascular display. *Anesth Analg* 2007 Nov;105(5):1303-1311 [doi: [10.1213/01.ane.0000282823.76059.ca](https://doi.org/10.1213/01.ane.0000282823.76059.ca)] [Medline: [17959959](https://pubmed.ncbi.nlm.nih.gov/17959959/)]
141. Anders S, Albert R, Miller A, Weinger MB, Doig AK, Behrens M, et al. Evaluation of an integrated graphical display to promote acute change detection in ICU patients. *Int J Med Inform* 2012 Dec;81(12):842-851 [FREE Full text] [doi: [10.1016/j.ijmedinf.2012.04.004](https://doi.org/10.1016/j.ijmedinf.2012.04.004)] [Medline: [22534099](https://pubmed.ncbi.nlm.nih.gov/22534099/)]
142. Effken JA, Loeb RG, Kang Y, Lin ZC. Clinical information displays to improve ICU outcomes. *Int J Med Inform* 2008 Nov;77(11):765-777 [doi: [10.1016/j.ijmedinf.2008.05.004](https://doi.org/10.1016/j.ijmedinf.2008.05.004)] [Medline: [18639487](https://pubmed.ncbi.nlm.nih.gov/18639487/)]
143. Faiola A, Newlon C. Advancing critical care in the ICU: a human-centered biomedical data visualization systems. In: *Proceedings of the 2011 International Conference on Ergonomics and Health Aspects of Work with Computers*. 2011 Presented at: EHAWC '11; July 9-14, 2011; Orlando, FL p. 119-128 URL: [https://link.springer.com/chapter/10.1007/978-3-642-21716-6\\_13](https://link.springer.com/chapter/10.1007/978-3-642-21716-6_13) [doi: [10.1007/978-3-642-21716-6\\_13](https://doi.org/10.1007/978-3-642-21716-6_13)]
144. Faiola A, Srinivas P, Duke J. Supporting clinical cognition: a human-centered approach to a novel ICU information visualization dashboard. *AMIA Annu Symp Proc* 2015;2015:560-569 [FREE Full text] [Medline: [26958190](https://pubmed.ncbi.nlm.nih.gov/26958190/)]
145. Forsman J, Anani N, Eghdam A, Falkenhav M, Koch S. Integrated information visualization to support decision making for use of antibiotics in intensive care: design and usability evaluation. *Inform Health Soc Care* 2013 Dec;38(4):330-353 [doi: [10.3109/17538157.2013.812649](https://doi.org/10.3109/17538157.2013.812649)] [Medline: [23957739](https://pubmed.ncbi.nlm.nih.gov/23957739/)]
146. Görges M, Kück K, Koch SH, Agutter J, Westenskow DR. A far-view intensive care unit monitoring display enables faster triage. *Dimens Crit Care Nurs* 2011;30(4):206-217 [doi: [10.1097/DCC.0b013e31821b7f08](https://doi.org/10.1097/DCC.0b013e31821b7f08)] [Medline: [21654229](https://pubmed.ncbi.nlm.nih.gov/21654229/)]
147. Wachter SB, Johnson K, Albert R, Syroid N, Drews F, Westenskow D. The evaluation of a pulmonary display to detect adverse respiratory events using high resolution human simulator. *J Am Med Inform Assoc* 2006 Nov;13(6):635-642 [FREE Full text] [doi: [10.1197/jamia.M2123](https://doi.org/10.1197/jamia.M2123)] [Medline: [16929038](https://pubmed.ncbi.nlm.nih.gov/16929038/)]
148. van Amsterdam K, Cnossen F, Ballast A, Struys M. Visual metaphors on anaesthesia monitors do not improve anaesthetists' performance in the operating theatre. *Br J Anaesth* 2013 May;110(5):816-822 [FREE Full text] [doi: [10.1093/bja/aes516](https://doi.org/10.1093/bja/aes516)] [Medline: [23384736](https://pubmed.ncbi.nlm.nih.gov/23384736/)]
149. Ordóñez P, Oates T, Lombardi ME, Hernandez G, Holmes K, Fackler J, et al. Visualization of multivariate time-series data in a neonatal ICU. *IBM J Res Dev* 2012 Sep;56(5):7:1-712 [FREE Full text] [doi: [10.1147/JRD.2012.2200431](https://doi.org/10.1147/JRD.2012.2200431)]
150. McKinlay A, McVittie C, Reiter E, Freer Y, Sykes C, Logie R. Design issues for socially intelligent user interfaces. A discourse analysis of a data-to-text system for summarizing clinical data. *Methods Inf Med* 2010;49(4):379-387 [doi: [10.3414/ME0613](https://doi.org/10.3414/ME0613)] [Medline: [20027380](https://pubmed.ncbi.nlm.nih.gov/20027380/)]
151. Jadhav A, Baldwin T, Wu J, Mukherjee V, Syeda-Mahmood T. Semantic expansion of clinician generated data preferences for automatic patient data summarization. *AMIA Annu Symp Proc* 2021;2021:571-580 [FREE Full text] [Medline: [35308964](https://pubmed.ncbi.nlm.nih.gov/35308964/)]
152. Laxmisan A, McCoy A, Wright A, Sittig D. Clinical summarization capabilities of commercially-available and internally-developed electronic health records. *Appl Clin Inform* 2017 Dec 16;03(01):80-93 [FREE Full text] [doi: [10.1055/s-0037-1618556](https://doi.org/10.1055/s-0037-1618556)]
153. Lin C. ROUGE: a package for automatic evaluation of summaries. *Text Summarization Branches Out*. URL: <https://aclanthology.org/W04-1013/> [accessed 2022-11-15]
154. Papineni K, Roukos S, Ward T, Zhu WJ. BLEU: a method for automatic evaluation of machine translation. In: *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*. 2002 Presented at: ACL '02; July 7-12, 2002; Philadelphia, PA p. 311-318 URL: <https://dl.acm.org/doi/10.3115/1073083.1073135> [doi: [10.3115/1073083.1073135](https://doi.org/10.3115/1073083.1073135)]
155. Rimmer A. Radiologist shortage leaves patient care at risk, warns royal college. *BMJ* 2017 Oct 11;359:j4683 [doi: [10.1136/bmj.j4683](https://doi.org/10.1136/bmj.j4683)] [Medline: [29021184](https://pubmed.ncbi.nlm.nih.gov/29021184/)]

156. Zhu X, Cimino JJ. Clinicians' evaluation of computer-assisted medication summarization of electronic medical records. *Comput Biol Med* 2015 Apr;59:221-231 [FREE Full text] [doi: [10.1016/j.combiomed.2013.12.006](https://doi.org/10.1016/j.combiomed.2013.12.006)] [Medline: [24393492](#)]
157. Pivovarov R, Coppleson YJ, Gorman SL, Vawdrey DK, Elhadad N. Can patient record summarization support quality metric abstraction? *AMIA Annu Symp Proc* 2016;2016:1020-1029 [FREE Full text] [Medline: [28269899](#)]
158. Van Vleck TT, Stein DM, Stetson PD, Johnson SB. Assessing data relevance for automated generation of a clinical summary. *AMIA Annu Symp Proc* 2007 Oct 11;2007:761-765 [FREE Full text] [Medline: [18693939](#)]
159. El-Kassas WS, Salama CR, Rafea AA, Mohamed HK. Automatic text summarization: a comprehensive survey. *Expert Syst Appl* 2021 Mar;165:113679 [doi: [10.1016/j.eswa.2020.113679](https://doi.org/10.1016/j.eswa.2020.113679)]
160. Tan J, Wan X, Xiao J. From neural sentence summarization to headline generation: a coarse-to-fine approach. In: *Proceedings of the 26th International Joint Conference on Artificial Intelligence*. 2017 Presented at: IJCAI'17; August 19-25, 2017; Melbourne, Australia p. 4109-4115 URL: <https://dl.acm.org/doi/abs/10.5555/3171837.3171860> [doi: [10.24963/ijcai.2017/574](https://doi.org/10.24963/ijcai.2017/574)]
161. Hou L, Hu P, Bei C. Abstractive document summarization via neural model with joint attention. In: *Proceedings of the 2017 National CCF conference on natural language processing and Chinese computing*. 2017 Presented at: NLPCC '17; November 8-12, 2017; Dalian, China p. 329-338 URL: [https://link.springer.com/chapter/10.1007/978-3-319-73618-1\\_28](https://link.springer.com/chapter/10.1007/978-3-319-73618-1_28) [doi: [10.1007/978-3-319-73618-1\\_28](https://doi.org/10.1007/978-3-319-73618-1_28)]
162. Moratanch N, Chitrakala S. A survey on extractive text summarization. In: *Proceedings of the 2017 International Conference on Computer, Communication and Signal Processing*. 2017 Presented at: ICCSP '17; January 10-11, 2017; Chennai, India p. 1-6 URL: <https://ieeexplore.ieee.org/document/7944061> [doi: [10.1109/iccsp.2017.7944061](https://doi.org/10.1109/iccsp.2017.7944061)]
163. Gupta V, Lehal GS. A survey of text summarization extractive techniques. *J Emerg Technol Web Intell* 2010 Aug 20;2(3):258-268 [FREE Full text] [doi: [10.4304/jetwi.2.3.258-268](https://doi.org/10.4304/jetwi.2.3.258-268)]
164. Savova G, Masanz J, Ogren P, Zheng J, Sohn S, Kipper-Schuler K, et al. Mayo clinical text analysis and knowledge extraction system (cTAKES): architecture, component evaluation and applications. *J Am Med Inform Assoc* 2010;17(5):507-513 [FREE Full text] [doi: [10.1136/jamia.2009.001560](https://doi.org/10.1136/jamia.2009.001560)] [Medline: [20819853](#)]
165. Alipour S, Eslami B, Abedi M, Ahmadinejad N, Arabkheradmand A, Aryan A, et al. A practical, clinical user-friendly format for breast ultrasound report. *Eur J Breast Health* 2021 Apr 1;17(2):165-172 [FREE Full text] [doi: [10.4274/ejbh.galenos.2021.6344](https://doi.org/10.4274/ejbh.galenos.2021.6344)] [Medline: [33870117](#)]
166. de Baca ME, Arnaut R, Brodsky V, Birdsong GG. Ordo ab Chao: framework for an integrated disease report. *Arch Pathol Lab Med* 2015 Feb;139(2):165-170 [FREE Full text] [doi: [10.5858/arpa.2013-0561-CP](https://doi.org/10.5858/arpa.2013-0561-CP)] [Medline: [25611099](#)]
167. Kay S. The international patient summary and the summarization requirement. *Stud Health Technol Inform* 2021 Oct 27;285:17-30 [doi: [10.3233/SHTI210569](https://doi.org/10.3233/SHTI210569)] [Medline: [34734848](#)]
168. Charette RS, Sarpong NO, Weiner TR, Shah RP, Cooper HJ. What's in a summary? Laying the groundwork for advances in hospital-course summarization. In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2021 Presented at: NAACL '21; June 6-11, 2021; Virtual Event p. 4794-4811 URL: <https://aclanthology.org/2021.naacl-main.382.pdf> [doi: [10.52198/22.sti.41.os1633](https://doi.org/10.52198/22.sti.41.os1633)]
169. Law AS, Freer Y, Hunter J, Logie RH, McIntosh N, Quinn J. A comparison of graphical and textual presentations of time series data to support medical decision making in the neonatal intensive care unit. *J Clin Monit Comput* 2005 Jun;19(3):183-194 [doi: [10.1007/s10877-005-0879-3](https://doi.org/10.1007/s10877-005-0879-3)] [Medline: [16244840](#)]
170. Bauer DT, Guerlain S, Brown PJ. The design and evaluation of a graphical display for laboratory data. *J Am Med Inform Assoc* 2010 Jul 01;17(4):416-424 [FREE Full text] [doi: [10.1136/jamia.2009.000505](https://doi.org/10.1136/jamia.2009.000505)] [Medline: [20595309](#)]
171. Salmon P, Rappaport A, Bainbridge M, Hayes G, Williams J. Taking the problem oriented medical record forward. *Proc AMIA Annu Fall Symp* 1996:463-467 [FREE Full text] [Medline: [8947709](#)]
172. Zhou L, Hripcsak G. Temporal reasoning with medical data--a review with emphasis on medical natural language processing. *J Biomed Inform* 2007 Apr;40(2):183-202 [FREE Full text] [doi: [10.1016/j.jbi.2006.12.009](https://doi.org/10.1016/j.jbi.2006.12.009)] [Medline: [17317332](#)]
173. Dziadzko MA, Herasevich V, Sen A, Pickering BW, Knight AA, Moreno Franco P. User perception and experience of the introduction of a novel critical care patient viewer in the ICU setting. *Int J Med Inform* 2016 Apr;88:86-91 [doi: [10.1016/j.jimedinf.2016.01.011](https://doi.org/10.1016/j.jimedinf.2016.01.011)] [Medline: [26878767](#)]
174. Li W, McCallum A. Pachinko allocation: DAG-structured mixture models of topic correlations. In: *Proceedings of the 23rd international conference on Machine learning*. 2006 Presented at: ICML '06; June 25-29, 2006; Pittsburgh, PA p. 577-584 URL: <https://dl.acm.org/doi/10.1145/1143844.1143917> [doi: [10.1145/1143844.1143917](https://doi.org/10.1145/1143844.1143917)]
175. Plaisant C, Mushlin R, Snyder A, Li J, Heller D, Shneiderman B. LifeLines: using visualization to enhance navigation and analysis of patient records. *Craft Inf Vis* 2003;308-312 [FREE Full text] [doi: [10.1016/b978-155860915-0/50038-x](https://doi.org/10.1016/b978-155860915-0/50038-x)]
176. Goldstein A, Shahar Y. Generation of natural-language textual summaries from longitudinal clinical records. In: Sarkar IN, Georgiou A, de Azevedo Marques PM, editors. *Studies in Health Technology and Informatics*. Amsterdam, Netherlands. ISO Press; 2015:594
177. Ghosh A. On the challenges of using evidence-based information: the role of clinical uncertainty. *J Lab Clin Med* 2004 Aug;144(2):60-64 [doi: [10.1016/j.lab.2004.05.013](https://doi.org/10.1016/j.lab.2004.05.013)] [Medline: [15322499](#)]
178. Chang H. Evaluation framework for telemedicine using the logical framework approach and a fishbone diagram. *Healthc Inform Res* 2015 Oct;21(4):230-238 [FREE Full text] [doi: [10.4258/hir.2015.21.4.230](https://doi.org/10.4258/hir.2015.21.4.230)] [Medline: [26618028](#)]

179. Lin YL, Guerguerian A, Laussen P. Heuristic evaluation of data integration and visualization software used for continuous monitoring to support intensive care: a bedside nurse's perspective. *J Nurs Care* 2015;04(06):1-8 [FREE Full text] [doi: [10.4172/2167-1168.1000300](https://doi.org/10.4172/2167-1168.1000300)]
180. Hsueh PS, Zhu XX, Hsiao MJ, Lee SY, Deng V, Ramakrishnan S. Automatic summarization of risk factors preceding disease progression an insight-driven healthcare service case study on using medical records of diabetic patients. *World Wide Web* 2014 Sep 14;18(4):1163-1175 [FREE Full text] [doi: [10.1007/s11280-014-0304-2](https://doi.org/10.1007/s11280-014-0304-2)]
181. Chaudhary A, George M, Chacko A. Extractive summarization of EHR notes. In: *Proceedings of the 2020 International Conference on Paradigms of Computing, Communication and Data Sciences*. 2020 Presented at: PCCDS '20; May 1-3, 2020; Kurukshetra, India p. 909-919 URL: [https://link.springer.com/chapter/10.1007/978-981-15-7533-4\\_73](https://link.springer.com/chapter/10.1007/978-981-15-7533-4_73) [doi: [10.1007/978-981-15-7533-4\\_73](https://doi.org/10.1007/978-981-15-7533-4_73)]
182. Anokwa Y, Ribeka N, Parikh T, Borriello G, Were M. Design of a phone-based clinical decision support system for resource-limited settings. In: *Proceedings of the 5th International Conference on Information and Communication Technologies and Development*. 2012 Presented at: ICTD '12; March 12-15, 2012; Atlanta, GA p. 13-24 URL: <https://dl.acm.org/doi/10.1145/2160673.2160676> [doi: [10.1145/2160673.2160676](https://doi.org/10.1145/2160673.2160676)]

## Abbreviations

**EHR:** electronic health record

**HCP:** health care professional

**ICU:** intensive care unit

**LIME:** Local Interpretable Model-Agnostic Explanations

**NLP:** natural language processing

**PRISMA-ScR:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

**ROUGE:** Recall-orientated Understudy for Gisting Evaluation

**UMLS:** Unified Medical Language System

*Edited by A Benis; submitted 28.11.22; peer-reviewed by R Perotte, PF Chen, E Sezgin, X Yan; comments to author 26.12.22; revised version received 15.03.23; accepted 25.07.23; published 28.11.23*

*Please cite as:*

*Keszthelyi D, Gaudet-Blavignac C, Bjelogrljic M, Lovis C*

*Patient Information Summarization in Clinical Settings: Scoping Review*

*JMIR Med Inform* 2023;11:e44639

URL: <https://medinform.jmir.org/2023/1/e44639>

doi: [10.2196/44639](https://doi.org/10.2196/44639)

PMID: [38015588](https://pubmed.ncbi.nlm.nih.gov/38015588/)

©Daniel Keszthelyi, Christophe Gaudet-Blavignac, Mina Bjelogrljic, Christian Lovis. Originally published in *JMIR Medical Informatics* (<https://medinform.jmir.org>), 28.11.2023. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Informatics*, is properly cited. The complete bibliographic information, a link to the original publication on <https://medinform.jmir.org/>, as well as this copyright and license information must be included.