

Review

# Empowering Health Care Education Through Learning Analytics: In-depth Scoping Review

Iva Bojic<sup>1\*</sup>, PhD; Maleyka Mammadova<sup>1\*</sup>, BA; Chin-Siang Ang<sup>1</sup>, PhD; Wei Lung Teo<sup>1</sup>, BA; Cristina Diordieva<sup>1</sup>, PhD; Anita Pienkowska<sup>1</sup>, PhD; Dragan Gašević<sup>2</sup>, PhD; Josip Car<sup>1</sup>, PhD

<sup>1</sup>Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

<sup>2</sup>Department of Human Centred Computing, Faculty of Information Technology, Monash University, Melbourne, Australia

\*these authors contributed equally

**Corresponding Author:**

Iva Bojic, PhD

Lee Kong Chian School of Medicine

Nanyang Technological University

11 Mandalay Rd

Singapore, 308232

Singapore

Phone: 65 65138572

Email: [iva.bojic@ntu.edu.sg](mailto:iva.bojic@ntu.edu.sg)

## Abstract

**Background:** Digital education has expanded since the COVID-19 pandemic began. A substantial amount of recent data on how students learn has become available for learning analytics (LA). LA denotes the “measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs.”

**Objective:** This scoping review aimed to examine the use of LA in health care professions education and propose a framework for the LA life cycle.

**Methods:** We performed a comprehensive literature search of 10 databases: MEDLINE, Embase, Web of Science, ERIC, Cochrane Library, PsycINFO, CINAHL, ICTP, Scopus, and IEEE Explore. In total, 6 reviewers worked in pairs and performed title, abstract, and full-text screening. We resolved disagreements on study selection by consensus and discussion with other reviewers. We included papers if they met the following criteria: papers on health care professions education, papers on digital education, and papers that collected LA data from any type of digital education platform.

**Results:** We retrieved 1238 papers, of which 65 met the inclusion criteria. From those papers, we extracted some typical characteristics of the LA process and proposed a framework for the LA life cycle, including digital education content creation, data collection, data analytics, and the purposes of LA. Assignment materials were the most popular type of digital education content (47/65, 72%), whereas the most commonly collected data types were the number of connections to the learning materials (53/65, 82%). Descriptive statistics was mostly used in data analytics in 89% (58/65) of studies. Finally, among the purposes for LA, understanding learners’ interactions with the digital education platform was cited most often in 86% (56/65) of papers and understanding the relationship between interactions and student performance was cited in 63% (41/65) of papers. Far less common were the purposes of optimizing learning: the provision of at-risk intervention, feedback, and adaptive learning was found in 11, 5, and 3 papers, respectively.

**Conclusions:** We identified gaps for each of the 4 components of the LA life cycle, with the lack of an iterative approach while designing courses for health care professions being the most prevalent. We identified only 1 instance in which the authors used knowledge from a previous course to improve the next course. Only 2 studies reported that LA was used to detect at-risk students during the course’s run, compared with the overwhelming majority of other studies in which data analysis was performed only after the course was completed.

(*J Med Internet Res* 2023;25:e41671) doi: [10.2196/41671](https://doi.org/10.2196/41671)

**KEYWORDS**

distance education and web-based learning; distributed learning environments; data science applications in education; 21st century abilities; cooperative and collaborative learning; COVID-19; education; digital; data; student

## Introduction

### Background

Health care education includes preregistration education, postregistration education, and lifelong learning for practicing clinicians [1]. The goal of health care professions education is to improve the potential of people—doctors, nurses, and allied health care professionals—to meet the health needs of patients, communities, and populations [2]. This requires ensuring that students learn what they need to know and be able to do, assessing the extent to which educational goals and outcomes are achieved and monitoring the quality of educational programs [1].

Digital education can be defined as learning through the use of electronic technology. It is becoming increasingly popular in academic and clinical settings [3] for continuing education [4]. Although education has evolved from an offline to a web-based environment through internet access, the COVID-19 pandemic has made web-based learning the standard delivery method worldwide [5,6]. Thus, the *new normal* educational practices based on technology-enabled globalization, borderless teaching, and seamless learning are emerging [7]. The transition from an in-person learning format to a digital learning format brings new challenges [8] but also opens new opportunities for educators through the use of a digital education platform. Digital education platforms are web-based systems that automate the administration of learning activities, including delivery, tracking, and reporting activities [9]. They subsequently generate substantial amounts of data on student learning, which are available for further analysis with learning analytics (LA).

LA can be defined as the “measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” [10]. LA refers to the collection and analysis of data from digital education platforms. Data analysis is used to determine student engagement in various digital learning activities to improve student learning. Specifically, LA differs from traditional educational analyses in 3 ways. First, data sets are typically much larger because of the strong quantitative focus, allowing for a higher degree of confidence in the generalizability of the results [11]. Second, data collected through technology has a finer granularity of variables than those collected via self-report, interview, or observational data [12]. Finally, the data are usually longitudinal [11], meaning that the methods and processes used to collect the data can incorporate a temporal dimension into the analysis [13]. These features provide opportunities for new insights by allowing learners to record, analyze, and retrieve their learning processes and outcomes for future use [14].

Given that LA provides powerful mechanisms for understanding, interpreting, and shaping learning processes and experiences using data-driven methods, research in this field has gained popularity since its inception as a scientific discipline in 2011

[15]. The number of studies using the keyword *learning analytics* in SciVerse Scopus between 2011 and 2021 was initially in the 2-digit range (n=29) but has since steadily increased, peaking at 917 in 2019. Despite a minor decline in the past 2 years (n=911 in 2020 and n=825 in 2021), the area continues to generate sizable interest.

### Literature Gaps and Research Questions

Although many studies have been conducted in the last decade that demonstrate the usefulness of LA for health care professions education, we have been able to identify only 2 reviews on this topic [14,16]. Their limitations were that both covered literature reviews only up to 2017, and one of the reviews included data from a single database (ie, MEDLINE). Moreover, they did not propose any framework to analyze the reviewed literature. On the basis of our literature review, which included papers from 10 databases, we identified 4 components of LA. Those 4 components were conceptualized in the form of a life cycle. The life cycle is a cyclical process that explains, in 4 stages, what digital learning content is created, what data are gathered from the content, and how the data are analyzed and used to achieve learning purposes. The first stage of the life cycle involves the creation of learning content. The content could be an assessment or some other type of digital material. The gathering of data occurs in the second stage. It entails gathering data from various sources, such as learners’ interactions with digital learning content and their performance on assessments and tests. This can be as simple as mouse clicks or key presses or as complex as time spent on different pages or interaction with videos or audio files. The interaction with the material generates data that can be analyzed using analytics tools in the third stage of the life cycle. Finally, in the fourth stage of the life cycle, the data analysis results would be used to help improve learner performance, the quality of learning experience, or both.

The idea of an LA life cycle is not new and was previously discussed in a study by Clow [17]. The study proposed a 4-step LA model that includes learners, data, metrics, and interventions. On the basis of the model, learners are people who learn, data are what we collect about learner behavior, metrics are numbers or statistics that describe data, and interventions are specific programs or sets of steps to provide learners with support in an area of need. Although our proposed life cycle model shared some similarities with Clow’s model, the main difference between the 2 models is that the first stage of our LA life cycle is referred to as *content* and not *learner*. We choose to exclude *learner* from our life cycle because previous research has shown that LA is highly content dependent [18]. When designing a course, it is critical to spell out content curation (the totality of what is to be taught in a course). It forms the basis for teaching by determining which topics, concepts, facts, skills, and values are expected to be taught and learned. With a deep dive into what goes into teaching and learning, we can make an informed decision about what variables and metrics can be tracked and analyzed to reveal learners’ preferences and engagement. In

other words, content has implications for what data are collected, how such data need to be preprocessed, what approaches can be applied to analyze data, and what can be done to optimize learning outcomes. The research questions that guided this review were as follows:

1. What types of content are used in LA studies in health care professions education?
2. What types of data are collected in health care professions education through content engagement?
3. What approaches are used for data analytics in health care professions education?
4. What are the purposes of analyzing the collected data in health care professions education?

## Methods

To provide a comprehensive overview of the use of LA in health care professions education, we analyzed the literature by conducting a scoping review. This study used the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) standards [19].

### Data Sources and Searches

We searched the following databases for papers written in English from January 2010 to October 2021: MEDLINE, Embase, Web of Science, ERIC, Cochrane Library, PsycINFO, CINAHL, ICTP, Scopus, and IEEE Explore. The year 2010 was used as the starting point for our search, as the field of LA was formally recognized after the first International Conference

on LA and Knowledge in 2011 [20]. We developed an initial search strategy based on a past publication on digital health education [21]. After initial title and abstract screening, we used the Search Refiner and Word Frequency Analyzer tools [22] to refine our search strategy further. Finally, we incorporated terms denoting different digital education platforms seen in the initially screened literature (Table S1 in [Multimedia Appendix 1](#)). Our final search strategy comprised (1) terms denoting health care professionals, (2) terms denoting different web-based learning environments, and (3) terms relating to LA. [Multimedia Appendix 2](#) contains the final search strategies for all databases.

### Study Selection

We included peer-reviewed publications that met the following criteria: papers on health care professions education, papers on digital education, and papers that collected LA data. We excluded papers that collected data about patient education; educators and not learners; and multimodal LA, abstracts, presentations, reviews, and commentary columns and papers where full texts could not be obtained ([Textbox 1](#)).

In total, 2 reviewers (IB and MM) performed title, abstract, and full-text screening of all publications retrieved using the first version of our search strategy. After the first round of screening, they defined a detailed set of inclusion and exclusion criteria and clearly explained the defined criteria to the other 4 reviewers (AP, CD, WLT, and CSA). All 6 reviewers worked in pairs sequentially and evaluated all retrieved papers using our final search strategy. We resolved disagreements on study selection by consensus and discussion with other reviewers.

**Textbox 1.** Inclusion and exclusion criteria for study selection.

<p><b>Inclusion criteria</b></p> <ul style="list-style-type: none"> <li>• Articles about health professions education</li> <li>• Articles about digital learning</li> <li>• Collected digital learning analytics data from a learning management system</li> </ul> <p><b>Exclusion criteria</b></p> <ul style="list-style-type: none"> <li>• Articles about patient education</li> <li>• Collected data about educators and not learners</li> <li>• Collected multimodal learning analytics data</li> <li>• Reviews and commentary columns</li> </ul>
---

### Data Extraction and Synthesis

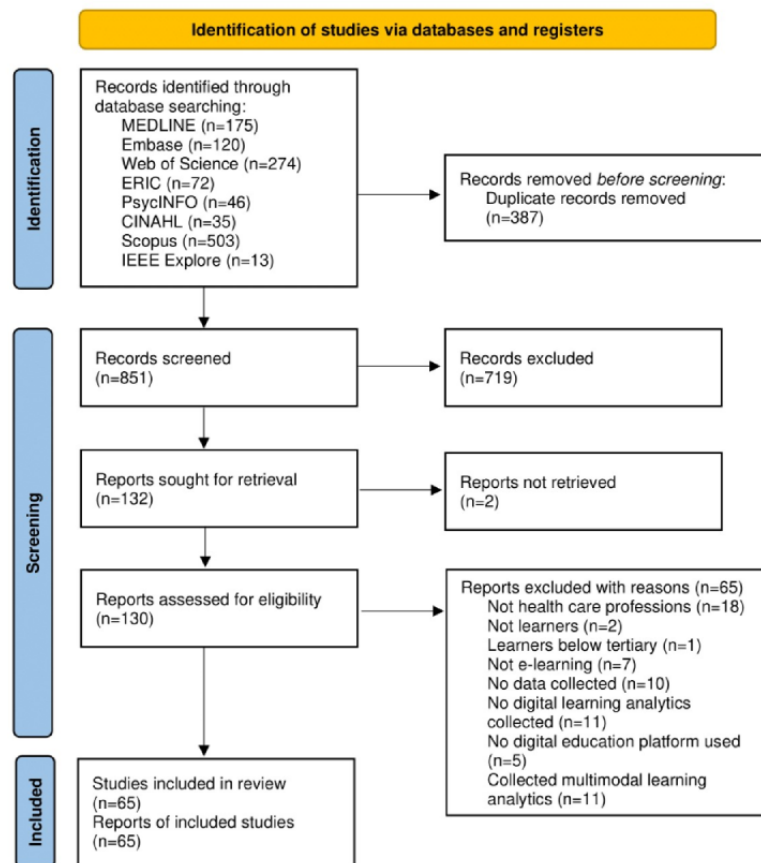
A total of 3 reviewers (IB, MM, and CSA) collaborated on a data charting form to select the variables to be extracted. In an iterative process, the 3 reviewers separately charted the data, reviewed the findings, and modified the data extraction form. The 3 reviewers then grouped study data by type of digital education content, type of data obtained from student interactions with digital education platforms, analytics approaches used to process the collected data, and finally the purposes of the LA. We resolved disagreements on data extraction by consensus and discussion with other reviewers. Upon the completion of data extraction, the data were collated, summarized, and subsequently organized into groups according

to similarities and differences. Tentative codes were added to these groups and iteratively discussed and revised to generate a comprehensive map of terms. The map was then used to conduct a narrative synthesis, which allowed for the identification of key themes and patterns within the data.

## Results

The search yielded 1238 papers ([Figure 1](#)), of which 851 were obtained after deduplication; 719 were excluded after title and abstract screening because they did not meet the inclusion criteria, and another 2 were excluded because of the unavailability of full texts. Full texts of 130 papers were retrieved and screened, of which 65 met the inclusion criteria.

**Figure 1.** PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) diagram of the literature search process.



### Study Characteristics

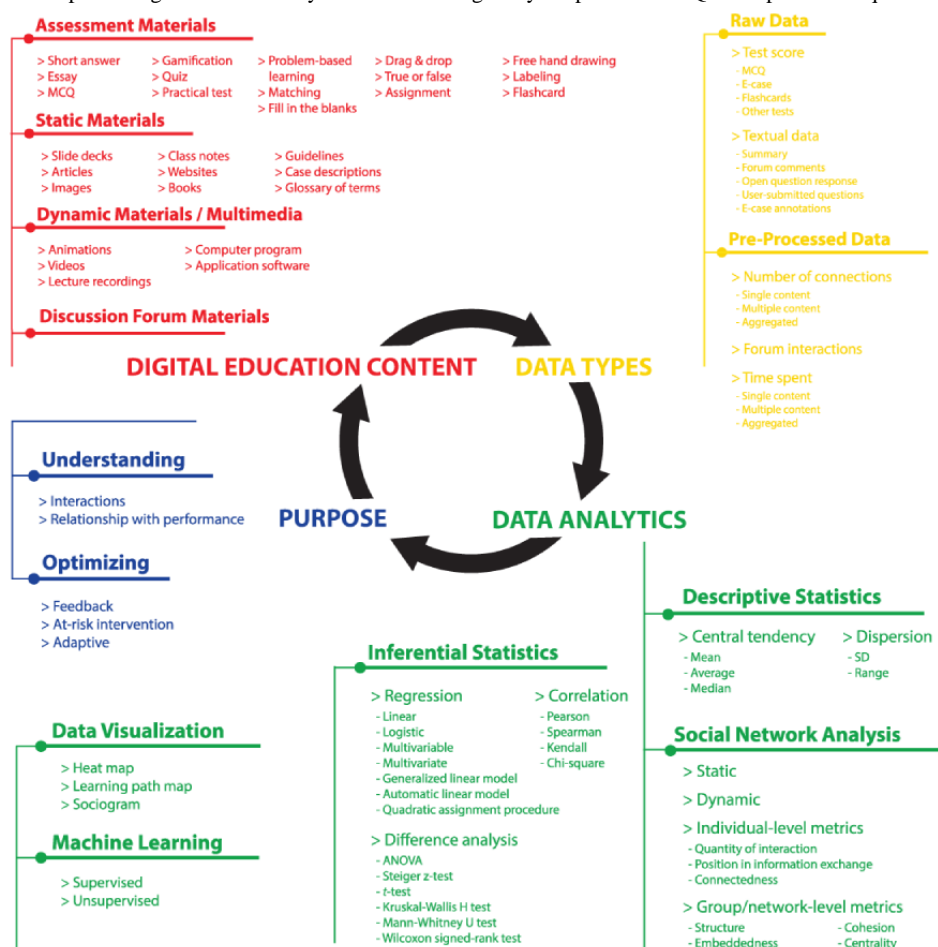
Among the included papers, the number of articles showed an upward inclination over the last decade, reaching a relative maximum in 2020 (n=14) and declining in 2021 (n=7). The most common study locations were the United States (n=12), Australia (n=11), Saudi Arabia (n=10), and Spain (n=6). Other countries had ≤5 papers, and 2 papers did not report the study location. A total of 60 studies were conducted in a university context with undergraduate and postgraduate students, and the remaining studies recruited health care professionals. The most common subject areas were medicine (n=32), dentistry (n=8), and nursing (n=7). Moodle [23] was the most commonly used digital education platform among the included articles (n=29), followed by Blackboard (n=4) [24,25]. A summary of article characteristics is shown in Table S2 in Multimedia Appendix 1.

### Life Cycle of LA

In our literature analysis, we identified several components of the LA process. We propose that these components can be conceptualized as part of a larger LA life cycle. Figure 2 shows the life cycle of LA, starting from digital education content

creation, followed by the data collection process, the use of analytics to process the collected data, and finally the purposes for which LA was used. Digital education content is mostly created by educators to support student learning, but in certain cases, it can also be created by learners’ peers (eg, discussion forum materials). While students are interacting with the provided digital education content, data on their use are collected and stored for further processing. The raw data include the click-level data that contain time-stamped information about every interaction with the digital education platform. The preprocessed data refer to the data that are mostly aggregated and shown in a dashboard. Both types of data can be used later in the process of analyzing the data. We identified 5 types of analytical approaches: (1) descriptive statistics, (2) inferential statistics, (3) data visualization, (4) machine learning, and (5) social network analysis. Finally, as stated in the definition of LA itself [10], the purpose of LA is to “understand and optimize learning and the environments in which it occurs.” Once the entire process of collecting and analyzing data is complete, the knowledge gained can be used to improve the produced digital education content, which then starts a new round of the entire LA life cycle.

**Figure 2.** The conceptual map showing the whole life cycle of the learning analytics process. MCQ: multiple-choice question.



### Research Question 1: What Types of Content Are Used in LA Studies in Health Care Professions Education?

Education content is any material that can be used to facilitate the learning process. Currently, many types of content are available for various learning purposes. Exploring the various types of content used in health care professions education is essential for understanding the contextual complexity that shapes the way we can build and use LA. The contents of the various studies examined in this review can be classified into 4 broad categories: (1) static materials, (2) dynamic materials or multimedia, (3) discussion forum materials, and (4) assessment materials (Table S3 in [Multimedia Appendix 1](#)). A total of 62% (40/65) of studies reported 1 or 2 types of digital education content [3,24,26-63], 28% (18/65) of studies reported 3 types [25,64-80], and the remaining 11% (7/65) of studies reported all 4 types [81-87].

*Static materials* are any type of digital education content that does not contain interactive elements. These include articles [81], guidelines [25], class notes [64], websites [26], case descriptions [27], glossaries of terms [28], images [65], slide decks [66], and books [82]. A total of 45% (29/65) of papers discussed providing static materials and outlining what specific materials were available for learners to view or read [25-29,42,58,60,61,63-70,72-74,76-78,80-82,84-86]. Another 9% (6/65) of studies used static materials but did not specify what materials were prepared for learners [36,44,75,79,83,87].

Among the various options, class notes appeared to be the most popular static material. Static materials remained popular because they can be easily viewed on any browser or mobile device and do not require additional plug-ins or technology [29].

*Dynamic materials* or *multimedia* are types of content that combine visual (eg, images, text, and animations) and auditory materials (eg, audio and podcast) in a single content package to convey a message [30]. Typically, they consist of animations [68], videos [69], and lecture recordings [31]. These materials were found to be the second most popular format and were discussed in 49% (32/65) of papers [24-26,30-35,46,47,52,56,57,64-71,74,76-78,80-82,84-86]. A total of 8% (5/65) of other studies did not specify which dynamic materials they had prepared for their learners [41,44,79,83,87]. Although videos remained the primary form of instruction in health care professions education, some studies have highlighted the importance of experiential learning [32]. A few studies have examined a computer program [33] or application software [71] with interactive controls and activities that provide learners with some level of control to create a sense of authenticity, immersion, and interaction. This would allow learners to contextualize theoretical knowledge in a simulated environment by linking new experiences to prior knowledge. This type of content is an excellent solution for training in health care professions in a safe environment without fear of distressing the patient [34]. Undeniably, dynamic materials or multimedia

are more versatile and promise greater engagement in learning than static text or images alone [35].

*Discussion forum materials* are provided to conduct specific learning activities outside the classroom, which are structured by educators [36]. The discussion forum is one of the most important social elements in a digital education course [3]. It is an effective tool to transform the classroom into a networked learning environment [86]. Students can actively engage in their learning process and develop important communication skills while posting, commenting, liking, and collaborating with other students on course content [83]. In total, 35% (23/65) of papers presented the use of a discussion forum to support student learning, although the materials used in the papers were not specified [3,36-38,48-51,54,59,62,64,71-73,75,81-87]. Understandably, each forum has its own set of category-based threads that are visible to all the learners in a course. These threads or posts create an active learning environment in which students regularly reflect on course content and actively collaborate to build knowledge through discussion. As discussion forums are asynchronous, they provide a safe environment for students to share ideas and learn from each other [73].

*Assessment materials* were the most popular digital education content, with 72% (47/65) of papers discussing them [24,25,27-30,32-35,39-43,45,46,52,53,55,56,58,60,61,65-87]. Assessment materials are important for improving the overall quality of learning. Learners' perceptions of when, how, and what they are assessed on affect how they learn. Assessment materials include drag and drop, fill in the blanks, true or false [81], matching, freehand drawing, labeling [68,74], flash cards [35,40], quizzes [75], multiple-choice questions (MCQs) [84], problem-based learning [85], practical tests [39], assignments [84], short answers [86], and essays [41]. Most assessment materials were administered at the end of a module or course so that educators could determine how much students have learned and retained and whether learning objectives have been met [4,25]. Of the 47 papers, 32 (68%) papers indicated that their assessment materials included feedback to help learners gain a better understanding of a topic and provide them with a reflection prompt [25,27,28,30,32-34,42,43,53,55,58,60,61,65-68,70-72,74,76,77,79-81,83-87]. Furthermore, assessment materials were also a fun and interactive way to create natural breaks between main topics and make learners feel like the course is progressing [65]. Although gamified assessments are possible, only 1 study has investigated this [43]. Students were found to be more motivated in the assessments when such features were effectively integrated.

## Research Question 2: What Types of Data Are Collected in Health Care Professions Education Through Content Engagement?

Data collected from the digital education platform can be broadly divided into 2 categories: raw and preprocessed data. Raw data consist of unprocessed data taken directly from the platform often in the form of log files. Many platforms automatically generate log files as a record of all user interactions with the system. For instance, the Moodle platform maintains a log of all user actions on the system and the time

they were made (Figure 3). Preprocessed data can then be obtained by processing the raw data and aggregating data points to focus on a specific aspect of it. For example, the same log data can be used to determine how often a student viewed a given course by extracting each instance of a specific user accessing that course. We found that most preprocessed data types fit into the broader categories of (1) *number of connections*, (2) *time spent*, and (3) *forum interactions* (Table S4 in Multimedia Appendix 1). Moreover, not all raw data consist of log files. Other types of raw data can be collected directly from relevant learning materials on the digital education platform, which we categorized into (4) *test scores* and (5) *textual data*.

*A number of connections* refers to data that measures the number of times users connect to the digital education platform. This is one of the most common data types collected in LA research and can be derived from log data from many platforms. Of articles that collected data on the number of connections to the platform, 32% (21/65) of collected data on multiple content types individually [24,26,27,29,36,37,39,56,58,63-65,67,73,76,78-80,82,85,86]. This allows for more detailed results by comparing the use of different digital education content types with student outcomes. Furthermore, 12% (8/65) of papers counted only the total, aggregated number of connections to the platform, without further dividing by content type [30,44,45,52,74,75,77,83], whereas 71% (25/65) of articles focused on connections to just 1 content type, such as in a study by Corrias and Cho [46], which measured student access to only the web-based video content [25,28,31,33,35,38,40,42,43,46-51,54,55,57,59,61,62,68-70,72]. A total of 26% (17/65) of papers used log data to generate timeline analytics, deriving insights into the distribution of connections over a day, week, semester, or even school year [31,33,39-43,47,53,55,57,65,68,70,71,75,86]. In so doing, researchers could determine the times at which materials were most frequently accessed and identify potential patterns, such as the relationship between material use and examination dates or the material access behaviors of higher-performing students.

*Time spent* measures the amount of time that users spend on the digital education platform. One way in which these data can be used is to understand how students interact with learning materials. For example, Cirigliano et al [65] measured the number of times students spent learning on each series of learning cards. They were able to distinguish between 3 types of student card interactions: those where students spent <20 seconds per card (rushing), those where they spent >100 seconds per card (idling), and those where they spent 20 to 100 seconds. Gaining a better understanding of rushing and idling can provide insights into whether students truly engage in digital education content. As with the number of connections, only 3% (2/65) of papers specified the amount of time spent on each content type [65,71], whereas 3% (2/65) of others calculated the overall time spent on the digital education platform [47,75]. The remaining (13/65, 20%) reviewed papers examined connections to only 1 content type [31,33,39-43,53,55,57,68,70,86].

*Forum interactions* consist of data derived from the students' use of web-based forums [3,37,38,48-51,54,59,62,64]. The data collected typically included the IDs of the sender and receiver,

the times that the post was written and modified, the post’s content, and the thread and forum where it was written. A total of 5% (3/65) of articles used forum interactions to observe how students engaged in problem-based learning that involved discussing concepts in web-based groups moderated by a tutor [48-50]. Forum interaction data were largely analyzed via social network analysis [51].

Test scores constitute data from assessments provided on a digital education platform. In most cases, researchers have collected the accuracy of participant performance in web-based clinical cases or e-cases [52,82]. The next largest group consisted of participants’ scores on the MCQs [41,43]. A total of 2% (1/65) of articles did not specify the type of assessment on which they collected data [75], and 2% (1/65) of articles featured not only MCQs but also fill in the blank, jumbled sentences, and matching questions [25]. In addition, 3% (2/65) of studies collected students’ performance on flash cards [35,40].

Textual data, such as written assessments, discussion forums, and email communications [88], can provide additional insights

into the learning process, which are not captured by quantitative data. Among the papers reviewed, we identified 17% (11/65) of papers that digitally collected textual data [3,24,25,34,41,60,66,76,79,84,86]. The majority (6/11, 55%) were open question responses from web-based surveys [24,25,60,76,79,86]. Other types of textual data collected included forum comments [3,84], user-submitted questions [66], e-case summary statements [41], and e-case annotations [34]. Textual data were generally analyzed manually, such as in a study by Colthorpe et al [24], where the authors manually categorized open question response data based on an existing classification system. However, 5% (3/65) of articles used machine learning tools to analyze textual data [34,41,66]. For example, Harrak et al [66] used an automatic annotation tool to analyze questions submitted by students to a digital education platform. The tool identified keywords in student questions, weighed them, and divided the questions into different sentences.

Figure 3. Example of a screenshot of a person’s log of actions performed on a Moodle course.

Time	User full name	Affected user	Event context	Component	Event name	Description	Origin	IP address
2021-12-27 12:11:32	Admin User	-	Course: Human Physiology	System	Course viewed	The user with id '1' viewed the course with id '101'.	web	0:0:0:0:0:0:1
2021-12-27 12:07:28	Admin User	-	Course: Human Physiology	System	Course viewed	The user with id '1' viewed the course with id '101'.	web	0:0:0:0:0:0:1
2021-12-27 11:50:24	Alice Ellis	-	Course: Human Physiology	System	Course viewed	The user with id '1002' viewed the course with id '101'.	web	0:0:0:0:0:0:1
2021-12-27 11:45:12	Alice Ellis	-	Course: Human Physiology	System	Course viewed	The user with id '1002' viewed the course with id '101'.	web	0:0:0:0:0:0:1
2021-12-27 11:44:15	Alice Ellis	-	Course: Human Physiology	System	Course viewed	The user with id '1002' viewed the course with id '101'.	web	0:0:0:0:0:0:1
2021-12-27 11:35:04	Alice Ellis	-	Book: Additional resources	Book	Chapter viewed	The user with id '1002' viewed the chapter with id '117' for the book with course module id '1091'.	web	0:0:0:0:0:0:1
2021-12-27 11:17:01	Alice Ellis	-	Book: Additional resources	Book	Chapter viewed	The user with id '1002' viewed the chapter with id '118' for the book with course module id '1091'.	web	0:0:0:0:0:0:1

### Research Question 3: What Approaches Are Used for Data Analytics in Health Care Professions Education?

We identified 5 analytic approaches used to process the collected data on students using digital education content (Table S5 in Multimedia Appendix 1): (1) *descriptive statistics*, (2) *inferential statistics*, (3) *data visualization*, (4) *machine learning*, and (5) *social network analysis*. Descriptive statistics summarize the characteristics of the collected data. Inferential statistics allow researchers to make predictions about the data [89]. Data visualization is used to present values visually [90]. Machine learning supports automated learning and makes predictions about collected data [91]. Social network analysis is the process of investigating social structures using the graph theory [76].

*Descriptive statistics* was the most commonly used approach to summarize the characteristics of the collected data: in 20% (13/65) of papers as the only approach [31,34,35,46,67,68,72-74,76,80,81,86] and in 89% (58/65) of papers as one of the used approaches [3,24-31,33-43,45-56,58-63,65-68,70,72-74,76-87]. Descriptive statistics can be further broken down into measures of *central tendency* (eg, average, mean, and median) and *dispersion* (eg, range and SD). The average is defined as the sum of all numbers divided by

the total number of values. The median is the middle number in a sorted list of numbers. Mean or average was more frequently used in the reviewed papers as a measure of central tendency than the median. The range is calculated as the maximum value minus the minimum value in the data set. SD measures the amount of data dispersion. Range and SD were almost equally used throughout the reviewed studies. Descriptive statistics measures were typically used to compare different student groups; for example, among the active, passive, and selective groups in a study by Ahmad et al [39]; English, Spanish, and French users in a study by Brands et al [81]; or different cohorts in a study by Alexander et al [67].

*Inferential statistics* used for hypothesis testing can be divided into *correlation*, *regression*, and *difference analysis*. This was the second most popular approach, used in 62% (40/65) of papers [24,25,29,30,33,36-41,43,45,47-56,58,59,62,63,65,66,70,71,75,77-79,82-85,87]. This correlation indicates whether the 2 data types collected are related. Regression investigates the relationship between an outcome (ie, dependent variables) and independent variables. The purpose of the difference analysis is to test the differences between the means of variables. Inferential statistics can be divided into *parametric* and *nonparametric* tests. Parametric tests make assumptions about

the parameters of the population distribution from which the sample is drawn, whereas nonparametric tests are distribution free.

A correlation coefficient (eg, Pearson, Kendall, and Spearman) is used to measure the strength of the relationship between numeric data types. For the Pearson correlation, both variables should be normally distributed. Kendall rank correlation and Spearman rank correlation do not carry any assumptions about the distribution of the data. Pearson correlation was used more often than Spearman correlation. Kendall rank correlation was only used in a study by Saqr et al [50]. Chi-square tests are used to determine whether nominal data types are associated. Chi-square tests were performed in a total of 15% (10/65) of papers [25,41,49,50,53,54,71,79,82,83]. The correlation was typically used to determine the variation in different data types collected through digital education platforms (eg, number of log-ins, total time, or time per question) and examination performance (eg, final grade) [29,55] or course completion [83].

Regression analysis can be divided into *linear* and *logistic* regression, depending on whether the outcome is continuous or dichotomous, or *multivariable* and *multivariate*, depending on whether there is 1 outcome or  $\geq 1$  outcome. Linear regression was used slightly more often than logistic regression. In addition to multivariate and multivariable regressions, generalized linear model was used in studies by Cirigliano et al [65] and Mu et al [53], automatic linear model was used in studies by Perumal et al [75] and Saqr et al [50], whereas the quadratic assignment procedure logistic regression was used in a study by Saqr and Montero [38]. Throughout the reviewed papers, the authors used collected data, such as age and attendance [25] and videos and quizzes accessed [56], as independent variables to predict dependent variables, such as student accuracy on the MCQ as a general indicator of successful engagement [65] or for the prediction of at-risk students [75].

Difference analysis or mean differences can be parametric (eg, *t* test or ANOVA) or nonparametric (eg, Mann-Whitney *U* test or Wilcoxon signed-rank test). In the papers, parametric tests were used in 22% (14/65) of instances [24,30,43,47,53-55,58,71,77,78,84,85,87], whereas nonparametric tests were used 14% (9/65) of the times [24,39,45,52,55,66,77,78,83]. Among parametric tests, ANOVA and *t* tests were used almost equally frequently. Mann-Whitney *U* test was more frequently used as a nonparametric test than others. For example, difference analysis was used to show substantial differences between an experimental group using a Smart Tutoring System in Moodle and a control group that did not use it [77] or showed that the students' use of digital education was associated with their subjects' final status (ie, success or failure) [78].

Data visualization techniques used in the analyzed papers can be divided into heat maps, learning path maps, and sociograms. Heat maps, used in 11% (7/65) of different studies, are visual representations of data that use color depth to make it more intuitive to interpret and evaluate the data [24,33,52,53,57,60,79]. Learning path maps, used to present the collected data in 11% (7/65) of papers, helped to visualize different course units that students master during a specific

subject or program [28,29,32,33,37,42,69]. A sociogram is a visual representation of a person's social connections. It was used in addition to other social network analysis techniques in 11% (7/65) of studies [38,48-51,54,59].

*Machine learning* approaches consist of 2 types: *supervised* and *unsupervised*. Supervised learning makes use of labeled data sets to train or *supervise* algorithms so that they can accurately identify data or forecast outcomes. In the reviewed papers, supervised machine learning was mostly used for classification tasks; for example, decision tree analysis was used to understand the relationship between the initially selected diagnostic hypotheses while examining web-based patient's cases and the final submitted hypothesis [27], support vector machines was used to predict students' final results in blended learning courses using student access time series generated from Moodle log [44], and naive Bayes network was used to provide students with feedback on the quality of their asked questions [66]. Unsupervised learning analyzes and clusters unlabeled data sets that can be used to identify hidden patterns in the data. k-means was the most popular unsupervised algorithm used in 9% (6/65) of papers, mostly as a mechanism for the early detection of at-risk students [24,26,58,66,77,79].

*Social network analysis* can be divided into individual-level and group/network-level measures and can be either static or dynamic. Individual-level measures are calculated on the level of an individual (ie, a student). Group or network-level data represent a group of students or the whole class (ie, the network). Individual-level measures can be further divided into a quantity of interaction, position in information exchange, and connectedness. Group or network-level measures included centrality, structure, embeddedness, and cohesion.

The quantity of interaction denotes how many interactions a student had in total (ie, degree centrality). One can also separately examine a measure of popularity (ie, in-degree centrality) or a measure of outgoing interaction diversity (ie, out-degree centrality). Position in information exchange (eg, betweenness centrality, information centrality, and closeness centrality) defines the role of students in information exchange in comparison with other students. Connectedness (eg, eigenvector centrality and clustering coefficient) denotes the proportion of the theoretical number of connections that have been achieved. Individual-level measures were calculated in almost all papers, except in a study by Jan and Vlachopoulos [3], where the authors focused only on calculating group-level measures, which were used for community detection.

Centrality measures denote the distribution of centrality measures within the group or network. The structure of the graph includes node count (ie, the number of students in each network), edge count (ie, the number of interactions), the average distance between all pairs of nodes, and average degree (ie, the average number of edges per node). Embeddedness includes network density and reciprocity. Network density is the proportion of actual interactions to the greatest feasible ratio among all members. Reciprocity is the ratio of reciprocated edges when 2 individuals exchange their responses. Cohesion includes efficiency, vulnerability, and transitivity. Efficiency is a valuable indicator of a network's robustness and tolerance



to limited failures or the removal of a small number of members. Vulnerability is a measure of graph sensitivity to information flow disturbance. Transitivity is the tendency of 2 individuals who have a common connection and a structural network characteristic that appears as triangles of nodes. Group-level measures were calculated in all 14% (9/65) of papers that used social network analysis [3,38,48-51,54,59,62].

In the reviewed papers, forum interactions were used as a type of data for building social networks to examine the properties of community-based [3] and problem-based learning [49,51]. More specifically, students' positions, interactions, and relationships in a network were used to predict their final performance. The most important predictors were information centrality and eigenvector centrality [50]. It was also shown that larger groups are associated with the decreased performance of individual students, as a high group size led to a less cohesive group, with less information exchange among students [48]. On the other hand, the well-performing groups were characterized by active and reciprocal interactions among students and group cohesion measures (transitivity and reciprocity) [59].

#### Research Question 4: What Are the Purposes of Analyzing the Collected Data in Health Care Professions Education?

The purposes for which LA data were used in the literature were (1) *understanding* and (2) *optimizing* learning and the contexts in which it occurs (Table S6 in [Multimedia Appendix 1](#)). Under the category of understanding, the purpose of LA can be further divided into the subcategories of understanding user *interactions* with the digital education platform and understanding the relationship between those interactions and user *performance*. The category of optimizing can also be subdivided into the subcategories of providing *feedback*, *at-risk intervention*, and *adaptive learning*.

*Understanding interactions* involves conducting LA to understand how students use the digital education platform. This could be as simple as examining the mean number of connections to a given material or as complicated as modeling how students make decisions in an e-case task. The interactions subcategory comprised the largest group in the papers we reviewed, with 91% (59/65) of papers examining patterns in student use of the platform [3,24-47,49,50,52,53,55-58,60-62,64-74,76-87]. Although more than a half of the papers (36/59, 61%) had at least one other purpose of understanding aside from understanding the interactions [24,25,29,30,33,34,36-41,43,44,47,49,50,52,53,55,58,62,64-68,71,77-79,82-86], 39% (23/59) of the papers focused only on understanding the interactions [3,26-28,31,32,35,42,45,46,56,57,60,61,69,70,72-74,76,80,81,87]. For instance, Brands et al [81] used a combination of user statistics from the course digital education platform and student course evaluations to gauge student use of and satisfaction with a web-based course. On a larger scale, Kim and Kim [80] aggregated LA data from 36 medical schools in Korea to examine trends in student use of digital education content.

*Understanding performance* involves the investigation of the relationship between learners' interactions with the digital

education platform and their performance. Performance can be evaluated in terms of student knowledge, skills, and attitudes. By examining the relationship between LA data and performance variables, authors can make conclusions about whether the use of digital education content leads to better outcomes and, if so, which type of content is more beneficial to students than others. As the second-largest category, a total of 65% (42/65) of papers fell under this group [24,25,29,30,33,34,36-41,43,44,47-55,58,59,62-68,71,75,77-79,82-86]. For example, Ahmad et al [39] used data on students' patterns of digital education content use to group students according to their time management strategies. They then examined the relationship between these time management strategies and students' academic performance.

*Feedback*, under the category of *optimizing* learning, involves the use of LA data to directly provide information to students about their learning. All articles (6/65, 9%) identified in this category provided feedback based on student performance on e-tests [33,34,53,55,58,77]. The feedback itself consisted simply of whether the students' responses were correct, often accompanied by a further explanation of the correct answer. For example, Sáiz-Manzanares et al [58] provided learners with feedback on web-based patient cases with MCQs about theoretical knowledge as well as actions that should be taken to treat the patient. Once the learner selected an answer, they received feedback indicating whether it was correct as well as the appropriate actions that needed to be taken for the patient (eg, "You should call a cardiology consultant immediately") and an explanation of the correct answer. In addition, hints were provided to guide students in performing the correct action.

*At-risk intervention* aims to provide additional support to students determined to be *at risk* of, for example, failing a course. To determine a student's risk, educators must first construct a predictive model that describes the relationship between predictors (eg, scores on formative assessments) and course failure. We identified 20% (13/65) of papers where the authors stopped short of presenting a model that could be used to predict student performance in the future [34,37,38,44,50,54,55,58,59,66,72,75,79]. However, only 1 study has put this predictive model into practice by providing interventions, namely, in a study by Sáiz-Manzanares et al [79], the authors constructed a model of students' risk of dropping out of the course based on which course educators provided personalized tutoring to at-risk students.

*Adaptive learning* in the educational context involves a system in which learners' interactions with digital content determine, at least partly, the nature of future content delivered to them [92]. In this review, we found only 6% (4/65) of instances of adaptive learning [29,30,40,57]. As with papers in the feedback subcategory, most of these studies used test scores as the basis for determining subsequent digital content. In the study by Linden et al [30], when students answered questions incorrectly, they were provided with a set of explanatory content, after which they had the opportunity to answer the same question again before being shown the next topic. In a study by Liu et al [29], the authors implemented a more involved form of adaptive learning: learners took a diagnostic test at the start of the module based on which the digital education platform created a

personalized learning path and study plan for the student. Moreover, students could use the Test Me function to test themselves again, which would cause corresponding changes in their study plan. Finally, the platform mentioned in a study by Menon et al [40] used a combination of learner accuracy and confidence on test questions to space the repetition of test items. An exception to this trend could be seen in a study by Gilliland [57], where the authors used the frequency of student engagement with particular topics on the learning platform as a proxy for the perceived difficulty of said content. They then communicated these findings to educators so that more teaching time could be allocated to challenging topics.

## Discussion

### Principal Findings

#### *Content Used in LA Studies*

Digital education has grown to be a critical component of higher education [3]. In the training of health care professionals, this review discovered that a variety of content types are used to achieve different types of outcomes (ie, knowledge, skills, and attitudes) [93]. The content can be text or audio, static or dynamic and in the form of video or animation [32,58,73,78]. Adding to the complication, the same content can be used by several users at different times and under different conditions. These factors make the standardization of data collected from various contexts extremely difficult. This also leads to associated difficulties in creating LA from preprocessing to analysis. Similar findings have been documented in some previous studies, indicating how difficult it is to generalize predictive models across different courses because of the differences in course designs [18]. These findings suggest that relevant metadata about the pedagogical purpose of various types of content should be ideally captured in the future to support the development of LA that can be built, deployed, and interpreted across different contexts [94].

Another issue is the creation of valid and reliable measures for various types of outcomes based on trace data used across different types of content [94,95]. Given the complexities of learning and teaching contexts, it is important to understand how different types of content can be used to support different pedagogical purposes [30,40]. Although current LA tools provide educators with some level of insight into how students interact with content, these tools do not provide educators with enough actionable insights to inform decisions about how to design personalized learning experiences for their students. As such, future research should focus on ways to improve our understanding of how learners interact with various types of content as well as develop better measures for various types of outcomes based on traced data.

In addition, it is worth noting that even when metadata about pedagogical intent are collected or when all the courses follow the same nominal pedagogy (eg, problem-based learning), developing a generalizable predictive model might not be so easy, as we often miss collecting data about individual differences among students. This has been shown in a recent paper with data from several years of problem-based learning

courses of a cohort of students enrolled in a medical school [96]. For example, some students are more likely to collaborate in group projects, whereas others prefer individual tasks; some students may spend more time reviewing their works before submission, whereas others submit their work immediately after completion; some students may tend to ask questions, whereas others do not; and some students are quick learners, whereas others require more efforts and time [14,46,80,82]. Therefore, it is critical to consider the individual differences among students when developing predictive or generalized models for LA.

#### *Data Collection*

LA, in its purest form, *passively* collects data from only 1 source. In contrast to LA, in multimodal LA, data are collected and integrated from different sources [97]. The purpose of multimodal LA is to *collect, synchronize, and analyze data from different communication modalities to provide on-time feedback* [98]. Using multimodal LA in health care professions education would allow for a more panoramic understanding of the learning processes and the different dimensions related to learning (ie, knowledge, skills, and attitudes). Moreover, although LA collects data about students seamlessly from digital education platforms, which makes it objective, it fails to capture individual differences and internal conditions of the learners (eg, their motivation, self-efficacy, cognitive load, or affective states) [99,100]. There is abundant evidence in the literature that showed the importance of motivational constructs as facilitators of academic self-regulation and achievement [101,102].

In the reviewed literature, there were a few studies in which the authors used multimodal LA to collect data while evaluating students' knowledge, skills, and attitudes. Health care professionals' *attitudes* toward learning were collected during their team-based learning session with videos and microphones in a study by Chua et al [103]; wearable sensors for heart rate, electrodermal activity, electroencephalography signals in studies by Antoniou et al [104,105]; and the Raspberry Pi minicomputer and Polar H10 chest belt in a study by Li et al [106]. Student *skills* were assessed by simulators in studies by Kennedy et al [107] and Chiu et al [108] while performing surgical tasks and by Microsoft Kinect in a study by Di Mitri et al [109] while performing cardiopulmonary resuscitation training. Finally, the gaze-following framework presented in a study by Barmaki and Guo [110] was used to track how *knowledge* of anatomy was gained in a collaborative effort. The aforementioned studies showed how multimodal LA can help collect data on all 3 components of learning separately. However, it is of vital importance to develop a comprehensive framework in which the learning of health care professions can be simultaneously tracked through all 3 prisms of learning, both in the digital and physical worlds.

A large amount of prior research has focused on self-report instruments used to measure self-regulation [111-113]. However, self-report instruments (eg, questionnaires, interviews, or learning diaries [114]) rely heavily on memories of the experience, where memories are biased and incomplete [115]. In addition, these instruments do not describe how learners dynamically adapt and modify their behavior during the actual learning process [116,117]. Previous research has shown that

trace-based data collected from digital education platforms can reflect self-reported data to a certain degree [118]. Although research on trace-based self-report measures is promising [99,119], the collection should be accompanied by incentives for students to self-report at the time when they complete their activities to avoid problems with conventional measures, which are usually not collected in real time [12]. An example of how this can be done is by collecting students' clicks on a 2D canvas [120]. In this sense, the 2D canvas would be used to report on self-efficacy and cognitive load for each activity in the course, while that click would also become a bookmark allowing students to easily retrieve activities that they felt unsure about or found difficult.

### Data Analytics

LA data are collected in real time, but data analytics is usually performed after course completion. The advantage of conducting real-time analytics is, for example, to implement early detection of at-risk students. Unfortunately, only a handful of studies have investigated how the detection of students at risk in combination with a personalized educational response to each student can be used to minimize students' dropout rates [58,79]. The second detected gap is that LA data are mostly collected over a period of up to 1 year or for 1 cohort of students (eg, 10 months in a study by Cirigliano et al [65], 1 academic year in a study by Koh et al [87], or a cohort of undergraduate students in a study by Kuhbeck et al [55]). Even when the data are collected for a longer period, they are mostly either jointly analyzed (eg, 3 years of data in a study by Brands et al [81]) or the comparison is made among different years (eg, 2 consecutive cohorts in a study by Chan et al [56]) or different courses (eg, 4 different courses in a study by Saqr et al [49]) without any changes made between 2 consecutive years or cohorts.

In cases where changes are implemented between 2 consecutive years or cohorts, the decision of what to change is very often not based on data analytics performed on the LA data previously collected. For example, in a study by Alvarez-Mendez et al [82], the authors introduced new tools, such as a glossary, a quiz, and a wiki, without relying on data analytics results. When comparing the performance of students before and after the aforementioned changes, Alvarez-Mendez et al [82] did not find significant relationships between the final grades and the digital education content provided. Among the reviewed papers, we identified only 1 paper [86] in which student feedback collected during the previous run was used as a tool for improving the quality of course content in subsequent years or cohorts. Thus, we propose that student feedback on digital education content or how teaching is conducted is used to improve the learning experience for subsequent years or cohorts. This would result in the closing of the LA life cycle loop.

Not only is LA offering valuable insights to educators on their students and courses, but it is also providing useful information to higher education institutions. This information could be potentially used to inform strategic decision-making regarding achieving educational excellence. However, we acknowledge that institutional planning and strategic decision-making processes are not easy tasks. This might be a consequence of a lack of understanding of (1) institutional culture within higher

education, (2) the degree to which individuals and cultures resist innovation and change, or (3) approaches to motivating social and cultural change [121,122]. Indeed, in many studies, it has been found that the most significant challenges that confront LA adoption are not technical but of a social nature [123-125]. In this context, and in the absence of a strategic goal or vision, LA data reporting has little power to start any significant changes.

### Purposes

Among the purposes for which LA can be used, most of the studies we reviewed used LA to understand learning, either to simply understand learners' interactions with the digital education platform or to determine the relationship between those interactions and learners' eventual performance. However, the measurement of performance tended to be limited to end-of-course examination results. Although these insights may be useful for course educators and administrators, they provided a limited picture of how these students would eventually perform as health care workers. One way to address this gap is to use other measures of performance, such as textual data and the aforementioned multimodal LA [126]. In addition, a longitudinal study that follows learners into clinical practice would provide uniquely pertinent insights into the effectiveness of digital education content in health care worker performance. This would allow the assessment of not only learners' theoretical knowledge but also their practical skills, as well as qualities such as communication with patients.

The second component, which was underrepresented in the reviewed studies, was the use of LA-derived insights to provide feedback, interventions, or adaptive learning. Of the 3 types of learning, feedback and interventions can both be considered as providing supplementary content to learners on a need basis, whereas adaptive learning can affect core course content as well by determining what modules a student is shown, and in which order, to suit the needs of the specific learner. However, we found few examples of researchers using LA-derived insights to fundamentally alter core course content or instructional design. One study of relevance was performed by Blakemore et al [84], who made iterative changes to a course's design over several runs to evaluate the efficacy of learning activities on student performance.

A similar dearth of research using LA to intervene in the learning environment has been noted in the general literature on LA [127]. The potential of harnessing LA to assess the effects of alterations in course content on student performance and engagement is a promising avenue explored in nonmedical areas [128]. Previous work has demonstrated the effectiveness of LA-based personalized feedback tools for improving such outcomes as learning effectiveness [129], student motivation [130], and academic performance [131]. The use of personalized feedback provision systems, such as *OnTask*, could present a promising and cost-effective avenue for health professions education to pursue in the future [132].

### Strengths and Limitations

Although this scoping review was conducted according to the scoping review methodology, some limitations are worth noting.

This review was guided by a formal protocol. To ensure a broad search of the literature, the search strategy included 10 databases; however, our inclusion criteria limited our review results to only English-language articles and published peer-reviewed literature from January 2010 to October 2021. Furthermore, as the data were extracted by 6 reviewers, we used the standardized extraction form to ensure consistency among reviewers. Nevertheless, the extraction process might be slightly affected by the presence of multiple reviewers. A final limitation is the rapid growth of the field; therefore, it is important to acknowledge that this scoping review is a snapshot at a particular point in time. Unlike traditional systematic reviews guided by well-defined constructs, in scoping reviews, it might be unfeasible to screen and synthesize all relevant literature on an emergent topic [133]. As our purpose was to primarily understand what LA techniques have been used in the education of health care professionals, our efforts to identify all eligible studies were limited in some respects.

### Practical Implications and Future Work

Our findings highlighted the need to develop an LA-specific implementation framework, drawing on empirical research related to LA efforts and drawing on existing knowledge and experience within the implementation science community. LA should be at the center of designing and redesigning courses for health care professionals. Namely, planning how LA should support the evolution of a course should come before the course starts rather than at its end. Finally, there is no need to repeat the mistakes that have already been made, and there is an opportunity to use the practical insights from existing research to provide an evidence-based foundation that can accelerate the

implementation of LA for the development of professional health courses.

### Conclusions

In this paper, we reviewed 65 papers in which the authors used LA to collect data on health care professions in their classes. From these papers, we extracted some general characteristics of the whole LA process and proposed a framework for the whole LA life cycle that is field independent, meaning it can be used beyond health care professions education. We then used the proposed framework to review the 65 papers.

Our study is student-centric, although we acknowledge that, in addition to students, there are other stakeholders such as educators and institutions involved in the entire LA process. We focused on students because most of the reviewed papers focused on students. This is consistent with the findings of Lee and Recker [133], where out of 47 reviewed papers, only 3 studies did not focus, at least partly, on students.

Finally, we identified the gaps in the current literature and proposed how they could be closed. The main limitation of almost all the papers we reviewed is that they did not use LA to iteratively inform course design and redesign. The purpose of collecting and analyzing LA data was mostly for an exploratory analysis or predictive analysis conducted within the borders of the same year or cohort. This means that the results of data analytics were not used to improve subsequent classes but rather to make some predictions about students who had already finished that class. In this sense, the potential of LA in health care professions education is not fully used.

### Acknowledgments

The authors would like to acknowledge the funding support from the World Health Organization collaborating centre for Digital Health and Health Education. JC's post at Imperial College London is supported by the National Institute for Health and Care Research Northwest London Applied Research Collaboration. Research of DG was in part supported by the Digital Health Cooperative Research Centre (DHCRC0056) and the Australian Research Council (DP210100060 and DP220101209).

### Data Availability

All relevant data are included in this publication. The titles of the papers included in this review are identified in the tables and [Multimedia Appendices 1](#) and [2](#). All excluded papers are available on request from the corresponding author. Researchers who require further details may contact the corresponding author with a methodologically sound proposal.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1

Tables listing extracted characteristics for each research question.

[\[XLSX File \(Microsoft Excel File\), 36 KB-Multimedia Appendix 1\]](#)

### Multimedia Appendix 2

Search strategies for all included databases.

[\[DOCX File , 56 KB-Multimedia Appendix 2\]](#)

### References

1. Transforming and scaling up health professionals' education and training. World Health Organization. 2019 Oct 01. URL: <https://www.who.int/publications/i/item/transforming-and-scaling-up-health-professionals%E2%80%99-education-and-training> [accessed 2023-04-07]
2. Institute of Medicine, Board on Health Care Services, Committee on the Health Professions Education Summit. Health Professions Education A Bridge to Quality. Washington, D.C., United States: National Academies Press; 2003.
3. Jan SK, Vlachopoulos P. Influence of learning design of the formation of online communities of learning. *Int Rev Res Open Distance Learn* 2018 Sep 26;19(4). [doi: [10.19173/irrodl.v19i4.3620](https://doi.org/10.19173/irrodl.v19i4.3620)]
4. Scott K, Morris A, Marais B. Medical student use of digital learning resources. *Clin Teach* 2018 Feb;15(1):29-33. [doi: [10.1111/tct.12630](https://doi.org/10.1111/tct.12630)] [Medline: [28300343](https://pubmed.ncbi.nlm.nih.gov/28300343/)]
5. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020 Feb;395(10223):497-506. [doi: [10.1016/s0140-6736\(20\)30183-5](https://doi.org/10.1016/s0140-6736(20)30183-5)]
6. Koh D. COVID-19 lockdowns throughout the world. *Occupational Med* 2020 May 02;70(5):322. [doi: [10.1093/occmed/kqaa073](https://doi.org/10.1093/occmed/kqaa073)]
7. Lockee BB. Online education in the post-COVID era. *Nat Electron* 2021 Jan 25;4(1):5-6. [doi: [10.1038/s41928-020-00534-0](https://doi.org/10.1038/s41928-020-00534-0)]
8. Pokhrel S, Chhetri R. A literature review on impact of COVID-19 pandemic on teaching and learning. *Higher Educ Future* 2021 Jan 19;8(1):133-141. [doi: [10.1177/2347631120983481](https://doi.org/10.1177/2347631120983481)]
9. Back DA, Behringer F, Haberstroh N, Ehlers JP, Sostmann K, Peters H. Learning management system and e-learning tools: an experience of medical students' usage and expectations. *Int J Med Educ* 2016 Aug 20;7:267-273 [FREE Full text] [doi: [10.5116/ijme.57a5.f0f5](https://doi.org/10.5116/ijme.57a5.f0f5)] [Medline: [27544782](https://pubmed.ncbi.nlm.nih.gov/27544782/)]
10. Gašević D, Dawson S, Siemens G. Let's not forget: learning analytics are about learning. *Tech Trends Tech Trends* 2014 Dec 21;59(1):64-71. [doi: [10.1007/s11528-014-0822-x](https://doi.org/10.1007/s11528-014-0822-x)]
11. Reimann P. Connecting learning analytics with learning research: the role of design-based research. *Learn Res Pract* 2016 Jul 22;2(2):130-142. [doi: [10.1080/23735082.2016.1210198](https://doi.org/10.1080/23735082.2016.1210198)]
12. Zhou M, Winne PH. Modeling academic achievement by self-reported versus traced goal orientation. *Learn Instruct* 2012 Dec;22(6):413-419. [doi: [10.1016/j.learninstruct.2012.03.004](https://doi.org/10.1016/j.learninstruct.2012.03.004)]
13. Saint J, Fan Y, Gašević D, Pardo A. Temporally-focused analytics of self-regulated learning: a systematic review of literature. *Comput Educ Artif Intell* 2022;3:100060. [doi: [10.1016/j.caeai.2022.100060](https://doi.org/10.1016/j.caeai.2022.100060)]
14. Saqr M. A literature review of empirical research on learning analytics in medical education. *Int J Health Sci (Qassim)* 2018;12(2):80-85 [FREE Full text] [Medline: [29599699](https://pubmed.ncbi.nlm.nih.gov/29599699/)]
15. Conole G, Gasevic D, Long P, Siemens G. Message from the LAK 2011 general and program chairs. In: Proceedings of the 1st International Conference on Learning Analytics and Knowledge Association for Computing Machinery (ACM). 2011 Presented at: 1st International Conference on Learning Analytics and Knowledge Association for Computing Machinery (ACM); Feb 27 -Mar 1, 2011; Banff, Canada.
16. Chan AK, Botelho MG, Lam OL. Use of learning analytics data in health care-related educational disciplines: systematic review. *J Med Internet Res* 2019 Feb 13;21(2):e11241 [FREE Full text] [doi: [10.2196/11241](https://doi.org/10.2196/11241)] [Medline: [30758291](https://pubmed.ncbi.nlm.nih.gov/30758291/)]
17. Clow D. The learning analytics cycle: closing the loop effectively. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. 2012 Presented at: LAK 2012: Second International Conference on Learning Analytics and Knowledge; Apr 29 - May 2, 2012; Vancouver, British Columbia, Canada. [doi: [10.1145/2330601.2330636](https://doi.org/10.1145/2330601.2330636)]
18. Gašević D, Dawson S, Rogers T, Gasevic D. Learning analytics should not promote one size fits all: the effects of instructional conditions in predicting academic success. *Internet Higher Educ* 2016 Jan;28:68-84. [doi: [10.1016/j.iheduc.2015.10.002](https://doi.org/10.1016/j.iheduc.2015.10.002)]
19. Peters M, Godfrey C, McInerney P, Soares C, Khalil H, Parker D. Methodology for JBI scoping reviews. In: The Joanna Briggs Institute Reviewers' Manual 2015. Adelaide, South Australia: Joanna Briggs Institute; Aug 2015.
20. Siemens G. Learning analytics: the emergence of a discipline. *Am Behav Sci* 2013 Aug 20;57(10):1380-1400. [doi: [10.1177/0002764213498851](https://doi.org/10.1177/0002764213498851)]
21. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital education in health professions: the need for overarching evidence synthesis. *J Med Internet Res* 2019 Feb 14;21(2):e12913 [FREE Full text] [doi: [10.2196/12913](https://doi.org/10.2196/12913)] [Medline: [30762583](https://pubmed.ncbi.nlm.nih.gov/30762583/)]
22. Clark J, Glasziou P, Del Mar C, Bannach-Brown A, Stehlik P, Scott AM. A full systematic review was completed in 2 weeks using automation tools: a case study. *J Clin Epidemiol* 2020 May;121:81-90. [doi: [10.1016/j.jclinepi.2020.01.008](https://doi.org/10.1016/j.jclinepi.2020.01.008)] [Medline: [32004673](https://pubmed.ncbi.nlm.nih.gov/32004673/)]
23. Rice W. Moodle 1.9 E-Learning Course Development: A complete guide to successful learning using Moodle 1st Edition. Birmingham: Packt Publishing; 2008.
24. Colthorpe K, Zimbardi K, Ainscough L, Anderson S. Know thy student! Combining learning analytics and critical reflections to develop a targeted intervention for promoting self-regulated learning. *Learn Analytics* 2015 May 29;2(1):134-155. [doi: [10.18608/jla.2015.21.7](https://doi.org/10.18608/jla.2015.21.7)]
25. Hughes M, Salamonson Y, Metcalfe L. Student engagement using multiple-attempt 'Weekly Participation Task' quizzes with undergraduate nursing students. *Nurse Educ Pract* 2020 Jul;46:102803. [doi: [10.1016/j.nepr.2020.102803](https://doi.org/10.1016/j.nepr.2020.102803)] [Medline: [32526682](https://pubmed.ncbi.nlm.nih.gov/32526682/)]

26. Judd T, Elliott K. Selection and use of online learning resources by first-year medical students: cross-sectional study. *JMIR Med Educ* 2017 Oct 02;3(2):e17 [[FREE Full text](#)] [doi: [10.2196/mededu.7382](https://doi.org/10.2196/mededu.7382)] [Medline: [28970187](https://pubmed.ncbi.nlm.nih.gov/28970187/)]
27. Doleck T, Basnet R, Poitras E, Lajoie S. Exploring the link between initial and final diagnosis in a medical intelligent tutoring system. In: Proceedings of the IEEE International Conference on MOOC, Innovation and Technology in Education (MITE). 2014 Presented at: IEEE International Conference on MOOC, Innovation and Technology in Education (MITE); Dec 19-20, 2014; Patiala, India. [doi: [10.1109/mite.2014.7020232](https://doi.org/10.1109/mite.2014.7020232)]
28. Doleck T, Basnet RB, Poitras EG, Lajoie SP. Mining learner–system interaction data: implications for modeling learner behaviors and improving overlay models. *J Comput Educ* 2015 Aug 5;2(4):421-447. [doi: [10.1007/s40692-015-0040-3](https://doi.org/10.1007/s40692-015-0040-3)]
29. Liu M, Kang J, Zou W, Lee H, Pan Z, Corliss S. Using data to understand how to better design adaptive learning. *Tech Know Learn* 2017 Jul 17;22(3):271-298. [doi: [10.1007/s10758-017-9326-z](https://doi.org/10.1007/s10758-017-9326-z)]
30. Linden K, Pemberton L, Webster L. Open Oceans: Learning without borders. Proceedings ASCILITE 2018. Victoria, Australia: Deakin University; 2018.
31. Garnett T, Button D. A case study exploring video access by students: wrangling and visualising data for measuring digital behaviour. In: Proceedings of the ASCILITE 2016. 2016 Presented at: Proceedings ASCILITE 2016; Nov 27-30, 2016; Adelaide, Australia.
32. Treasure-Jones T, Dent-Spargo R, Dharmaratne S. How do students want their workplace-based feedback visualized in order to support self-regulated learning? Initial results and reflections from a co-design study in medical education. In: Proceedings of the 13th European Conference On Technology Enhanced Learning. 2018 Presented at: 13th European Conference On Technology Enhanced Learning; Sep 3-6, 2018; Leeds, United Kingdom.
33. Pecaric M, Boutis K, Beckstead J, Pusic M. A big data and learning analytics approach to process-level feedback in cognitive simulations. *Acad Med* 2017;92(2):175-184. [doi: [10.1097/acm.0000000000001234](https://doi.org/10.1097/acm.0000000000001234)]
34. Poitras EG, Naismith L, Doleck T, Lajoie SP. Using learning analytics to identify medical student misconceptions in an online virtual patient environment. *Online Learn J* 2015 Dec 29;20(2). [doi: [10.24059/olj.v20i2.802](https://doi.org/10.24059/olj.v20i2.802)]
35. Tackett S, Green D, Dyal M, O'Keefe E, Thomas TE, Nguyen T, et al. Use of commercially produced medical education videos in a cardiovascular curriculum: multiple cohort study. *JMIR Med Educ* 2021 Oct 07;7(4):e27441 [[FREE Full text](#)] [doi: [10.2196/27441](https://doi.org/10.2196/27441)] [Medline: [34617911](https://pubmed.ncbi.nlm.nih.gov/34617911/)]
36. Broadbent J. Academic success is about self-efficacy rather than frequency of use of the learning management system. *Aust J Educ Technol* 2016 Jul 11;32(4). [doi: [10.14742/ajet.2634](https://doi.org/10.14742/ajet.2634)]
37. Saqr M, Nouri J, Fors U. Temporality matters. A learning analytics study of the patterns of interactions and its relation to performance. In: Proceedings of the 10th International Conference on Education and New Learning Technologies. 2018 Presented at: 10th International Conference on Education and New Learning Technologies; Jul 2-4, 2018; Palma, Spain. [doi: [10.21125/edulearn.2018.1305](https://doi.org/10.21125/edulearn.2018.1305)]
38. Saqr M, Montero C. Learning and social networks - similarities, differences and impact. In: Proceedings of the IEEE 20th International Conference on Advanced Learning Technologies (ICALT). 2020 Presented at: IEEE 20th International Conference on Advanced Learning Technologies (ICALT); Jul 06-09, 2020; Tartu, Estonia. [doi: [10.1109/ICALT49669.2020.00047](https://doi.org/10.1109/ICALT49669.2020.00047)]
39. Ahmad UN, Gasevic D, Matcha W, Jovanovic J, Pardo A, Lim L, et al. Discovering time management strategies in learning processes using process mining techniques. In: Transforming Learning with Meaningful Technologies. Cham: Springer; 2019.
40. Menon A, Gaglani S, Haynes MR, Tackett S. Using “big data” to guide implementation of a web and mobile adaptive learning platform for medical students. *Med Teach* 2017 May 16;1-6. [doi: [10.1080/0142159x.2017.1324949](https://doi.org/10.1080/0142159x.2017.1324949)]
41. Berman NB, Artino AR. Development and initial validation of an online engagement metric using virtual patients. *BMC Med Educ* 2018 Sep 17;18(1):213 [[FREE Full text](#)] [doi: [10.1186/s12909-018-1322-z](https://doi.org/10.1186/s12909-018-1322-z)] [Medline: [30223825](https://pubmed.ncbi.nlm.nih.gov/30223825/)]
42. Doleck T, Jarrell A, Poitras E, Chaouachi M, Lajoie S. Examining diagnosis paths: a process mining approach. In: Proceedings of the second International Conference on Computational Intelligence & Communication Technology (CICT). 2016 Presented at: Second International Conference on Computational Intelligence & Communication Technology (CICT); Feb 12-13, 2016; Ghaziabad, India. [doi: [10.1109/cict.2016.137](https://doi.org/10.1109/cict.2016.137)]
43. Roche CC, Wingo NP, Westfall AO, Azuero A, Dempsey DM, Willig JH. Educational analytics: a new frontier for gamification? *Comput Inform Nurs* 2018 Sep;36(9):458-465 [[FREE Full text](#)] [doi: [10.1097/CIN.0000000000000455](https://doi.org/10.1097/CIN.0000000000000455)] [Medline: [29985195](https://pubmed.ncbi.nlm.nih.gov/29985195/)]
44. Gamulin J, Gamulin O, Kermek D. Using Fourier coefficients in time series analysis for student performance prediction in blended learning environments. *Expert Syst* 2015 Dec 04;33(2):189-200. [doi: [10.1111/exsy.12142](https://doi.org/10.1111/exsy.12142)]
45. Mohseni Z, Martins R, Milrad M, Masiello I. Improving classification in imbalanced educational datasets using over-sampling. In: Proceedings of the 28th International Conference on Computers in Education. 2020 Presented at: 28th International Conference on Computers in Education; Nov 23-27, 2020; Online.
46. Corrias A, Cho HJ. Design and implementation of a flipped classroom learning environment in the biomedical engineering context. In: Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2015 Presented at: 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Aug 25-29, 2015; Milan, Italy. [doi: [10.1109/embc.2015.7319267](https://doi.org/10.1109/embc.2015.7319267)]

47. Choi-Lundberg DL, Cuellar WA, Williams AM. Online dissection audio-visual resources for human anatomy: undergraduate medical students' usage and learning outcomes. *Anat Sci Educ* 2016 Nov 24;9(6):545-554. [doi: [10.1002/ase.1607](https://doi.org/10.1002/ase.1607)] [Medline: [27802370](https://pubmed.ncbi.nlm.nih.gov/27802370/)]
48. Saqr M, Nouri J, Jormanainen I. A learning analytics study of the effect of group size on social dynamics and performance in online collaborative learning. In: *Transforming Learning With Meaningful Technologies*. Cham: Springer; 2019.
49. Saqr M, Fors U, Nouri J. Using social network analysis to understand online Problem-Based Learning and predict performance. *PLoS One* 2018 Sep 20;13(9):e0203590 [FREE Full text] [doi: [10.1371/journal.pone.0203590](https://doi.org/10.1371/journal.pone.0203590)] [Medline: [30235227](https://pubmed.ncbi.nlm.nih.gov/30235227/)]
50. Saqr M, Fors U, Tedre M. How the study of online collaborative learning can guide teachers and predict students' performance in a medical course. *BMC Med Educ* 2018 Feb 06;18(1):24 [FREE Full text] [doi: [10.1186/s12909-018-1126-1](https://doi.org/10.1186/s12909-018-1126-1)] [Medline: [29409481](https://pubmed.ncbi.nlm.nih.gov/29409481/)]
51. Saqr M, Alamro A. The role of social network analysis as a learning analytics tool in online problem based learning. *BMC Med Educ* 2019 May 22;19(1):160 [FREE Full text] [doi: [10.1186/s12909-019-1599-6](https://doi.org/10.1186/s12909-019-1599-6)] [Medline: [31113441](https://pubmed.ncbi.nlm.nih.gov/31113441/)]
52. Thau E, Perez M, Pusic MV, Pecaric M, Rizzuti D, Boutis K. Image interpretation: learning analytics-informed education opportunities. *AEM Educ Train* 2021 Apr 22;5(2):e10592 [FREE Full text] [doi: [10.1002/aet2.10592](https://doi.org/10.1002/aet2.10592)] [Medline: [33898916](https://pubmed.ncbi.nlm.nih.gov/33898916/)]
53. Mu EW, Pusic M, Coneybeare M, Meehan SA. Learning analytics to enhance dermatopathology education among dermatology residents. *J Drugs Dermatol* 2019 Dec 01;18(12):1231-1236. [Medline: [31860211](https://pubmed.ncbi.nlm.nih.gov/31860211/)]
54. Saqr M, Nouri J, Vartiainen H, Tedre M. Robustness and rich clubs in collaborative learning groups: a learning analytics study using network science. *Sci Rep* 2020 Sep 02;10(1):14445 [FREE Full text] [doi: [10.1038/s41598-020-71483-z](https://doi.org/10.1038/s41598-020-71483-z)] [Medline: [32879398](https://pubmed.ncbi.nlm.nih.gov/32879398/)]
55. Kühbeck F, Berberat PO, Engelhardt S, Sarikas A. Correlation of online assessment parameters with summative exam performance in undergraduate medical education of pharmacology: a prospective cohort study. *BMC Med Educ* 2019 Nov 08;19(1):412 [FREE Full text] [doi: [10.1186/s12909-019-1814-5](https://doi.org/10.1186/s12909-019-1814-5)] [Medline: [31703687](https://pubmed.ncbi.nlm.nih.gov/31703687/)]
56. Chan AK, Botelho MG, Lam OL. The relation of online learning analytics, approaches to learning and academic achievement in a clinical skills course. *Eur J Dent Educ* 2021 Aug 03;25(3):442-450. [doi: [10.1111/eje.12619](https://doi.org/10.1111/eje.12619)] [Medline: [33185309](https://pubmed.ncbi.nlm.nih.gov/33185309/)]
57. Gilliland KO. The flipped classroom and learning analytics in histology. *Med Sci Educ* 2017 Jan 12;27(1):9-13. [doi: [10.1007/s40670-016-0364-2](https://doi.org/10.1007/s40670-016-0364-2)]
58. Sáiz-Manzanas MC, Marticorena-Sánchez R, García-Osorio CI. Monitoring students at the university: design and application of a moodle plugin. *Applied Sci* 2020 May 18;10(10):3469. [doi: [10.3390/app10103469](https://doi.org/10.3390/app10103469)]
59. Saqr M, Nouri J, Vartiainen H, Malmberg J. What makes an online problem-based group successful? A learning analytics study using social network analysis. *BMC Med Educ* 2020 Mar 18;20(1):80 [FREE Full text] [doi: [10.1186/s12909-020-01997-7](https://doi.org/10.1186/s12909-020-01997-7)] [Medline: [32188471](https://pubmed.ncbi.nlm.nih.gov/32188471/)]
60. de Leng B, Pawelka F. The use of learning dashboards to support complex in-class pedagogical scenarios in medical training: how do they influence students' cognitive engagement? *Res Pract Technol Enhanced Learn* 2020 Dec;15(1). [doi: [10.1186/s41039-020-00135-7](https://doi.org/10.1186/s41039-020-00135-7)]
61. Lajoie SP, Poitras EG, Jarrell A, Doleck T. Modeling metacognitive activities in medical problem-solving with BioWorld. *Intell Syst Ref Library* 2015 Oct;76:323-343. [doi: [10.1007/978-3-319-11062-2\\_13](https://doi.org/10.1007/978-3-319-11062-2_13)]
62. Saqr M, Nouri J. High resolution temporal network analysis to understand and improve collaborative learning. In: *Proceedings of the 10th International Conference on Learning Analytics & Knowledge (LAK'20)*. 2020 Presented at: 10th International Conference on Learning Analytics & Knowledge (LAK'20); Mar 23–27, 2020; Frankfurt, Germany.
63. Vázquez PA, Álvarez-Méndez AM, Angulo-Carrere TA, Barrios JC. Learning analytics in human histology reveals different student' clusters and different academic performance. In: *Proceedings of the 14th International Technology, Education and Development Conference*. 2020 Presented at: 14th International Technology, Education and Development Conference; Mar 02-04, 2020; Valenica.
64. Penteadó B, Isotani S, Paiva P, Marina M, Ferrari D. Detecting behavioral trajectories in continued education online courses. In: *Proceedings of the IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*. 2019 Presented at: IEEE 19th International Conference on Advanced Learning Technologies (ICALT); Jul 15-18, 2019; Maceio, Brazil. [doi: [10.1109/icalt.2019.00069](https://doi.org/10.1109/icalt.2019.00069)]
65. Cirigliano MM, Guthrie CD, Pusic MV. Click-level learning analytics in an online medical education learning platform. *Teach Learn Med* 2020 May 12;32(4):410-421. [doi: [10.1080/10401334.2020.1754216](https://doi.org/10.1080/10401334.2020.1754216)] [Medline: [32397923](https://pubmed.ncbi.nlm.nih.gov/32397923/)]
66. Harrak F, Bouchet F, Luengo V. From students' questions to students' profiles in a blended learning environment. *Learn Analytics* 2019 Apr 14;6(1):54-84. [doi: [10.18608/jla.2019.61.4](https://doi.org/10.18608/jla.2019.61.4)]
67. Alexander J, Barcellona M, McLachlan S, Sackley C. Technology-enhanced learning in physiotherapy education: student satisfaction and knowledge acquisition of entry-level students in the United Kingdom. *Res Learn Technol* 2019 Feb 06;27. [doi: [10.25304/rlt.v27.2073](https://doi.org/10.25304/rlt.v27.2073)]
68. Perumal V, Butson R, Blyth P, Daniel B. Clinical anatomy e-cases: a five-year follow-up of learning analytics. *N Z Med J* 2017 Jan 27;130(1449):22-29. [Medline: [28178726](https://pubmed.ncbi.nlm.nih.gov/28178726/)]
69. Olmos M, Corrin L. Academic analytics in a medical curriculum: enabling educational excellence. *Australas J Educ Technol* 2012 Feb 01;28(1). [doi: [10.14742/ajet.880](https://doi.org/10.14742/ajet.880)]

70. Walldén S, Mäkinen E. Educational data mining and problem-based learning. *Inform Educ* 2014;13(1):141-156. [doi: [10.15388/infedu.2014.08](https://doi.org/10.15388/infedu.2014.08)]
71. Burbano GD, Soler J. Learning analytics in m-learning: periodontic education. In: *Telematics and Computing*. Cham: Springer; 2020.
72. Poitras E, Lajoie S, Doleck T, Jarrell A. Subgroup discovery with user interaction data: an empirically guided approach to improving intelligent tutoring systems. *Educ Technol Society* 2016;19(2):204-214.
73. Vazquez M, Barrios J, Granizo J. Focusing on undergraduates' behaviour. Learning analytics in human histology. In: *Proceedings of the 12th International Technology, Education and Development Conference*. 2018 Presented at: 12th International Technology, Education and Development Conference; Mar 5-7, 2018; Valencia, Spain. [doi: [10.21125/inted.2018.0516](https://doi.org/10.21125/inted.2018.0516)]
74. Perumal V, Daniel B, Butson R. Learning analytics of clinical anatomy e-cases. In: *Big Data and Learning Analytics in Higher Education*. Cham: Springer; 2017.
75. Saqr M, Fors U, Tedre M. How learning analytics can early predict under-achieving students in a blended medical education course. *Med Teach* 2017 Apr 19;39(7):757-767. [doi: [10.1080/0142159x.2017.1309376](https://doi.org/10.1080/0142159x.2017.1309376)]
76. Scott J. Social network analysis. *Sociology* 2016 Jul 02;22(1):109-127. [doi: [10.1177/0038038588022001007](https://doi.org/10.1177/0038038588022001007)]
77. Sáiz-Manzanares MC, Marticorena-Sánchez R, Díez-Pastor JF, García-Osorio CI. Does the use of learning management systems with hypermedia mean improved student learning outcomes? *Front Psychol* 2019 Feb 12;10:88 [FREE Full text] [doi: [10.3389/fpsyg.2019.00088](https://doi.org/10.3389/fpsyg.2019.00088)] [Medline: [30809162](https://pubmed.ncbi.nlm.nih.gov/30809162/)]
78. Queiroga EM, Enríquez CR, Cechinel C, Casas AP, Paragarino VR, Bencke LR, et al. Using virtual learning environment data for the development of institutional educational policies. *Applied Sci* 2021 Jul 24;11(15):6811. [doi: [10.3390/app11156811](https://doi.org/10.3390/app11156811)]
79. Sáiz-Manzanares MC, Rodríguez-Díez JJ, Díez-Pastor JF, Rodríguez-Arribas S, Marticorena-Sánchez R, Ji YP. Monitoring of student learning in learning management systems: an application of educational data mining techniques. *Applied Sci* 2021 Mar 17;11(6):2677. [doi: [10.3390/app11062677](https://doi.org/10.3390/app11062677)]
80. Kim K, Kim G. Development of e-learning in medical education: 10 years' experience of Korean medical schools. *Korean J Med Educ* 2019 Sep;31(3):205-214 [FREE Full text] [doi: [10.3946/kjme.2019.131](https://doi.org/10.3946/kjme.2019.131)] [Medline: [31455050](https://pubmed.ncbi.nlm.nih.gov/31455050/)]
81. Brands B, Tran NN, Baudendistel-Happ E, Sanchez-Garcia M, Fischer MR, Koletzko B. Global e-learning in early nutrition and lifestyle for international healthcare professionals: design and evaluation of the early nutrition specialist programme (ENS). *Nutrients* 2021 Feb 27;13(3):775 [FREE Full text] [doi: [10.3390/nu13030775](https://doi.org/10.3390/nu13030775)] [Medline: [33673535](https://pubmed.ncbi.nlm.nih.gov/33673535/)]
82. Alvarez-Mendez A, Carrere T, Barrios J, Bravo-Llatas C, Vazquez M. Application of data mining in moodle platform for the analysis of the academic performance of a compulsory subject in university students. In: *Proceedings of the 14th International Technology, Education and Development Conference*. 2020 Mar Presented at: 14th International Technology, Education and Development Conference; Mar 2-4 ,2020; Valencia, Spain.
83. Saeki YM, Murakami R, Tsuzuku S, Suzuki M, Nakano H, Suzuki K. Using learning analytics to build a learning support program for distance learning in nursing. In: *Proceedings of the 8th International Congress on Advanced Applied Informatics (IIAI-AAI)*. 2019 Presented at: 8th International Congress on Advanced Applied Informatics (IIAI-AAI); Jul 07-11 , 2019; Toyama, Japan. [doi: [10.1109/iiiai-aaai.2019.00076](https://doi.org/10.1109/iiiai-aaai.2019.00076)]
84. Blakemore LM, Meek SE, Marks LK. Equipping learners to evaluate online health care resources: longitudinal study of learning design strategies in a health care massive open online course. *J Med Internet Res* 2020 Feb 26;22(2):e15177 [FREE Full text] [doi: [10.2196/15177](https://doi.org/10.2196/15177)] [Medline: [32130120](https://pubmed.ncbi.nlm.nih.gov/32130120/)]
85. Chan AK, Botelho MG, Lam OL. An exploration of student access to a learning management system-challenges and recommendations for educators and researchers. *Eur J Dent Educ* 2021 Nov 03;25(4):846-855. [doi: [10.1111/eje.12664](https://doi.org/10.1111/eje.12664)] [Medline: [33470506](https://pubmed.ncbi.nlm.nih.gov/33470506/)]
86. King D, Arnaiz I, Gordon-Thomson C, Randal N, Herkes S. Evaluation and use of an online data acquisition and content platform for physiology practicals and tutorials. *Int J Innov Sci Math Educ* 2016;24(5):24-34.
87. Koh YY, Schmidt HG, Low-Beer N, Rotgans JI. Team-based learning analytics. *Acad Med* 2020;95(6):872-878. [doi: [10.1097/acm.0000000000003157](https://doi.org/10.1097/acm.0000000000003157)]
88. Lang C, Siemens G, Wise A, Gasevic D. *Handbook of Learning Analytics – First edition*. Beaumont, Alberta: Society for Learning Analytics Research; 2017.
89. Byrne G. A statistical primer: understanding descriptive and inferential statistics. *Evidence Based Library Inform Pract* 2007 Mar 14;2(1):32-47. [doi: [10.18438/B8FW2H](https://doi.org/10.18438/B8FW2H)]
90. Healy K. *Data Visualization A Practical Introduction*. Princeton, New Jersey, United States: Princeton University Press; 2018.
91. Mohri M, Rostamizadeh A, Talwalkar A. *Foundations of Machine Learning*. Cambridge, Massachusetts, United States: MIT press; 2018.
92. Kerr P. Adaptive learning. *ELT J* 2015 Oct 07;70(1):88-93. [doi: [10.1093/elt/ccv055](https://doi.org/10.1093/elt/ccv055)]
93. Epstein RM. Assessment in medical education. *N Engl J Med* 2007 Jan 25;356(4):387-396. [doi: [10.1056/NEJMra054784](https://doi.org/10.1056/NEJMra054784)] [Medline: [17251535](https://pubmed.ncbi.nlm.nih.gov/17251535/)]



94. Gašević D, Greiff S, Shaffer DW. Towards strengthening links between learning analytics and assessment: challenges and potentials of a promising new bond. *Comput Human Behav* 2022 Sep;134:107304. [doi: [10.1016/j.chb.2022.107304](https://doi.org/10.1016/j.chb.2022.107304)]
95. Winne PH. Construct and consequential validity for learning analytics based on trace data. *Comput Human Behav* 2020 Nov;112:106457. [doi: [10.1016/j.chb.2020.106457](https://doi.org/10.1016/j.chb.2020.106457)]
96. Jovanović J, Saqr M, Joksimović S, Gašević D. Students matter the most in learning analytics: the effects of internal and instructional conditions in predicting academic success. *Comput Educ* 2021 Oct;172:104251. [doi: [10.1016/j.compedu.2021.104251](https://doi.org/10.1016/j.compedu.2021.104251)]
97. Blikstein P, Worsley M. Multimodal learning analytics and education data mining: using computational technologies to measure complex learning tasks. *Learn Analytics* 2016 Sep 17;3(2):220-238. [doi: [10.18608/jla.2016.32.11](https://doi.org/10.18608/jla.2016.32.11)]
98. Echeverría V, Domínguez F, Chiluita K. Towards a distributed framework to analyze multimodal data. In: *Proceedings of the 1st International Workshop on Learning Analytics Across Physical and Digital Spaces, CrossLAK 2016*. 2016 Presented at: 1st International Workshop on Learning Analytics Across Physical and Digital Spaces, CrossLAK 2016; Apr 25-29, 2016; Edinburgh, Scotland. [doi: [10.1089/big.2021.0326](https://doi.org/10.1089/big.2021.0326)]
99. Siadaty M, Gasevic D, Hatala M. Trace-based micro-analytic measurement of self-regulated learning processes. *Learn Analytics* 2016 Apr 23;3(1):183-214. [doi: [10.18608/jla.2016.31.11](https://doi.org/10.18608/jla.2016.31.11)]
100. Srivastava N, Fan Y, Rakovic M, Singh S, Jovanovic J, Graaf J, et al. Effects of internal and external conditions on strategies of self-regulated learning: a learning analytics study. In: *Proceedings of the LAK22: 12th International Learning Analytics and Knowledge Conference*. 2022 Presented at: LAK22: 12th International Learning Analytics and Knowledge Conference; Mar 21 - 25, 2022; Online USA.
101. Zimmerman B, Schunk D. An essential dimension of self-regulated learning. In: *Motivation and Self-regulated Learning: Theory, Research, and Applications*. Hillsdale, NJ: Lawrence Erlbaum Associates Publishers; 2008:1-30.
102. Hsieh P, Sullivan JR, Sass DA, Guerra NS. Undergraduate engineering students' beliefs, coping strategies, and academic performance: an evaluation of theoretical models. *J Experiment Educ* 2012 Jan;80(2):196-218. [doi: [10.1080/00220973.2011.596853](https://doi.org/10.1080/00220973.2011.596853)]
103. Chua Y, Rajalingam P, Tan S, Dauwels J. EduBrowser: a multimodal automated monitoring system for co-located collaborative learning. In: *Learning Technology for Education Challenges*. Cham: Springer; 2019.
104. Antoniou P, Arfaras G, Pandria N, Ntakakis G, Bambatsikos E, Athanasiou A. Real-time affective measurements in medical education, using virtual and mixed reality. In: *Proceedings of the International Conference on Brain Function Assessment in Learning*. 2020 Presented at: International Conference on Brain Function Assessment in Learning; Oct 9-11, 2020; Heraklion, Crete, Greece. [doi: [10.1007/978-3-030-60735-7\\_9](https://doi.org/10.1007/978-3-030-60735-7_9)]
105. Antoniou PE, Arfaras G, Pandria N, Athanasiou A, Ntakakis G, Babatsikos E, et al. Biosensor real-time affective analytics in virtual and mixed reality medical education serious games: cohort study. *JMIR Serious Games* 2020 Sep 02;8(3):e17823 [FREE Full text] [doi: [10.2196/17823](https://doi.org/10.2196/17823)] [Medline: [32876575](https://pubmed.ncbi.nlm.nih.gov/32876575/)]
106. Li MX, Nadj M, Maedche A, Ifenthaler D, Wöhler J. Towards a physiological computing infrastructure for researching students' flow in remote learning. *Tech Know Learn* 2021 Sep 23;27(2):365-384. [doi: [10.1007/s10758-021-09569-4](https://doi.org/10.1007/s10758-021-09569-4)]
107. Kennedy G, Ioannou I, Zhou Y, Bailey J, O'Leary S. Mining interactions in immersive learning environments for real-time student feedback. *Australas J Educ Technol* 2013 May 10;29(2). [doi: [10.14742/ajet.700](https://doi.org/10.14742/ajet.700)]
108. Chiu H, Kang Y, Wang W, Chen C, Hsu W, Tseng M, et al. The role of active engagement of peer observation in the acquisition of surgical skills in virtual reality tasks for novices. *J Surg Educ* 2019 Nov;76(6):1655-1662. [doi: [10.1016/j.jsurg.2019.05.004](https://doi.org/10.1016/j.jsurg.2019.05.004)] [Medline: [31130508](https://pubmed.ncbi.nlm.nih.gov/31130508/)]
109. Di Mitri D, Schneider J, Specht M, Drachsler H. Detecting mistakes in CPR training with multimodal data and neural networks. *Sensors (Basel)* 2019 Jul 13;19(14):3099 [FREE Full text] [doi: [10.3390/s19143099](https://doi.org/10.3390/s19143099)] [Medline: [31337029](https://pubmed.ncbi.nlm.nih.gov/31337029/)]
110. Barmaki R, Guo Z. Deep neural networks for collaborative learning analytics: evaluating team collaborations using student gaze point prediction. *Australas J Educ Technol* 2020 Dec 28;36(6):53-71. [doi: [10.14742/ajet.6436](https://doi.org/10.14742/ajet.6436)]
111. Winne P, Perry N. Chapter 16 - measuring self-regulated learning. In: *Handbook of Self-Regulation*. Cambridge, Massachusetts, United States: Academic Press; 2000.
112. Hadwin AF, Nesbit JC, Jamieson-Noel D, Code J, Winne PH. Examining trace data to explore self-regulated learning. *Metacognition Learn* 2007 Nov 29;2(2-3):107-124. [doi: [10.1007/s11409-007-9016-7](https://doi.org/10.1007/s11409-007-9016-7)]
113. Winne PH. Improving measurements of self-regulated learning. *Educ Psychol* 2010 Oct 25;45(4):267-276. [doi: [10.1080/00461520.2010.517150](https://doi.org/10.1080/00461520.2010.517150)]
114. Roth A, Ogrin S, Schmitz B. Assessing self-regulated learning in higher education: a systematic literature review of self-report instruments. *Educ Asse Eval Acc* 2015 Sep 17;28(3):225-250. [doi: [10.1007/s11092-015-9229-2](https://doi.org/10.1007/s11092-015-9229-2)]
115. ElSayed A, Caeiro-Rodriguez M, MikicFonte F, Llamas-Nistal M. Research in learning analytics and educational data mining to measure self-regulated learning: a systematic review. In: *Proceedings of the 18th World Conference on Mobile and Contextual Learning*. 2019 Presented at: 18th World Conference on Mobile and Contextual Learning; Sep 16 - 18, 2019; Delft University of Technology, Netherlands.
116. Panadero E, Klug J, Järvelä S. Third wave of measurement in the self-regulated learning field: when measurement and intervention come hand in hand. *Scandinavian J Educ Res* 2015 Jul 24;60(6):723-735. [doi: [10.1080/00313831.2015.1066436](https://doi.org/10.1080/00313831.2015.1066436)]

117. Kovanović V, Gašević D, Joksimović S, Hatala M, Adesope O. Analytics of communities of inquiry: effects of learning technology use on cognitive presence in asynchronous online discussions. *Internet Higher Educ* 2015 Oct;27:74-89. [doi: [10.1016/j.iheduc.2015.06.002](https://doi.org/10.1016/j.iheduc.2015.06.002)]
118. Ye D, Pennisi S. Using trace data to enhance Students' self-regulation: a learning analytics perspective. *Internet Higher Educ* 2022 Jun;54:100855. [doi: [10.1016/j.iheduc.2022.100855](https://doi.org/10.1016/j.iheduc.2022.100855)]
119. You JW. Identifying significant indicators using LMS data to predict course achievement in online learning. *Internet Higher Educ* 2016 Apr;29:23-30. [doi: [10.1016/j.iheduc.2015.11.003](https://doi.org/10.1016/j.iheduc.2015.11.003)]
120. Jovanovic J, Gasevic D, Pardo A, Dawson S, Whitelock-Wainwright A. Introducing meaning to clicks: towards traced-measures of self-efficacy and cognitive load. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. 2019 Presented at: LAK19: The 9th International Learning Analytics & Knowledge Conference; Mar 4 - 8, 2019; Tempe, AZ, USA. [doi: [10.1145/3303772.3303782](https://doi.org/10.1145/3303772.3303782)]
121. Kavanagh MH, Ashkanasy NM. The impact of leadership and change management strategy on organizational culture and individual acceptance of change during a merger. *Br J Manage* 2006 Mar;17(S1):S81-103. [doi: [10.1111/j.1467-8551.2006.00480.x](https://doi.org/10.1111/j.1467-8551.2006.00480.x)]
122. Macfadyen L, Dawson S. Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *Educ Technol Society* 2012;15(3):149-163.
123. Howell JA, Roberts LD, Seaman K, Gibson DC. Are we on our way to becoming a “helicopter university”? Academics' views on learning analytics. *Tech Know Learn* 2017 Aug 2;23(1):1-20. [doi: [10.1007/s10758-017-9329-9](https://doi.org/10.1007/s10758-017-9329-9)]
124. Tsai Y, Moreno-Marcos PM, Jivet I, Scheffel M, Tammets K, Kollom K, et al. The SHEILA framework: informing institutional strategies and policy processes of learning analytics. *Learn Analytics* 2018 Nov 18;5(3):5-20. [doi: [10.18608/jla.2018.53.2](https://doi.org/10.18608/jla.2018.53.2)]
125. Tsai Y, Kovanović V, Gašević D. Connecting the dots: an exploratory study on learning analytics adoption factors, experience, and priorities. *Internet Higher Educ* 2021 Jun;50:100794. [doi: [10.1016/j.iheduc.2021.100794](https://doi.org/10.1016/j.iheduc.2021.100794)]
126. Chan T, Sebok-Syer S, Thoma B, Wise A, Sherbino J, Pusic M. Learning analytics in medical education assessment: the past, the present, and the future. *AEM Educ Train* 2018 Apr 22;2(2):178-187 [FREE Full text] [doi: [10.1002/aet2.10087](https://doi.org/10.1002/aet2.10087)] [Medline: [30051086](https://pubmed.ncbi.nlm.nih.gov/30051086/)]
127. Motz B, Bergner Y, Brooks C, Gladden A, Gray G, Lang C, et al. A Lak of direction: misalignment between the goals of learning analytics and its research scholarship. *J Learn Analytics* 2023 Mar 09:1-13. [doi: [10.18608/jla.2023.7913](https://doi.org/10.18608/jla.2023.7913)]
128. Hui B, Farvolden S. How can learning analytics improve a course? In: *Proceedings of the 22nd Western Canadian Conference on Computing Education*. 2017 Presented at: WCCCE '17: Western Canadian Conference on Computing Education; May 5 - 6, 2017; Abbotsford, BC, Canada.
129. Wang D, Han H. Applying learning analytics dashboards based on process - oriented feedback to improve students' learning effectiveness. *J Comput Assist Learn* 2020 Oct 26;37(2):487-499. [doi: [10.1111/jcal.12502](https://doi.org/10.1111/jcal.12502)]
130. Corrin L, Barba P. Exploring students' interpretation of feedback delivered through learning analytics dashboards. In: *Proceedings of the Australasian Society for Computers in Learning in Tertiary Education*. 2014 Presented at: Australasian Society for Computers in Learning in Tertiary Education; Nov 23 – 26, 2014; Dunedin, NZ.
131. Arnold K, Pistilli M. Course signals at Purdue: using learning analytics to increase student success. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. 2012 Presented at: LAK 2012: Second International Conference on Learning Analytics and Knowledge; Apr 29 - May 2, 2012; Vancouver, British Columbia, Canada.
132. Pardo A, Bartimote K, Buckingham Shum S, Dawson S, Gao J, Gašević D, et al. OnTask: delivering data-informed, personalized learning support actions. *J Learn Analytics* 2018 Dec 11;5(3):235-249. [doi: [10.18608/jla.2018.53.15](https://doi.org/10.18608/jla.2018.53.15)]
133. Lee J, Recker M. What do studies of learning analytics reveal about learning and instruction? A systematic literature review. In: *Learning, Design, and Technology*. Cham: Springer; 2018.

---

## Abbreviations

**LA:** learning analytics

**MCQ:** multiple-choice question

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

*Edited by T Leung; submitted 04.08.22; peer-reviewed by P Zhang, B Sousa-Pinto, R Erschens; comments to author 17.11.22; revised version received 28.11.22; accepted 08.03.23; published 17.05.23*

*Please cite as:*

*Bojic I, Mammadova M, Ang CS, Teo WL, Diordieva C, Pienkowska A, Gašević D, Car J  
Empowering Health Care Education Through Learning Analytics: In-depth Scoping Review  
J Med Internet Res 2023;25:e41671*

*URL: <https://www.jmir.org/2023/1/e41671>*

*doi: [10.2196/41671](https://doi.org/10.2196/41671)*

*PMID:*

©Iva Bojic, Maleyka Mammadova, Chin-Siang Ang, Wei Lung Teo, Cristina Diordieva, Anita Pienkowska, Dragan Gašević, Josip Car. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 17.05.2023. 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 the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.