Available online at **globets.org/journal**
International Journal of Education, Technology and Science
3(3) (2023) 427–453

IJETS
International Journal of Education Technology and Science

A SCOPING REVIEW OF THE "AT-RISK" STUDENT LITERATURE IN HIGHER EDUCATION

Colin Chibaya[a] *

[a] *Sol Plaatje University, South Africa*

Received: 22.05.2023 Revised version received: 19.07.2023 Accepted: 20.07.2023

Abstract

Higher Education Institutions have a strong desire to meet the goal of producing quality graduates. It is important for them to understand their student population to provide equal opportunities for diverse groups, including "at-risk" students. A scoping review of literature was conducted to understand the concept of "at-risk" students in Higher Education Institutions. The objectives of the study were to identify common categories of "at-risk" students, analyze the methods and variables used to study this topic, and understand the key factors and approaches used in this research. Literature was selected using the PCC framework and the JBI protocol and screened using the PRISMA-ScR framework. A total of 84 articles out of 1961 were eligible and included in the review. The results showed that there is a lack of research on "at-risk" students in Africa, but significant growth in related research in America, Europe, and Asia. Key findings include the importance of academic data and the use of statistical and machine learning methods. The factors that put students at risk are often linked to high school education, and ethnicity, gender, and location also play a role. Higher Education Institutions should implement interventions to address students' psychosocial well-being to create a supportive learning environment. This review is expected to provide insights for addressing similar challenges in the African context.

Keywords: "At-risk" student, dropout, stop-out, burn out, data analytics tools, intervention

*Corresponding author: Colin Chibaya. ORCID ID.: https://orcid.org/0000-0001-6995-605X
E-mail: colin.chibaya@spu.ac.za

1. Introduction

A scoping review on the topic of "at-risk" students in higher education can provide a comprehensive understanding of the current knowledge base and research landscape related to this population of students. It includes the definition and identification of "at-risk" students together with the factors contributing to their risk, and the interventions and programs being implemented to support them. The purpose of this scoping review is to identify gaps in the existing literature to inform future research directions, policy, and practice to better support students.

Higher Education Institutions pay close attention to research on "at-risk" students, over other vulnerable groups, for various reasons. Traditionally, Higher Education Institutions prioritize comprehensive understanding of their diverse student population, including "at-risk" students. They commonly aspire to improve completion rates, closing achievement gaps, preventing dropouts and failure, while fostering inclusive and supportive learning environments where equity and success are emphasized. To achieve these aims, Higher Education Institutions work to identify and address specific student needs, promoting equal educational opportunities and supporting academic and personal growth by reducing emotional stress. This study conducts a scoping review to understand the factors that distinguish "at-risk" students, to help Higher Education Institutions better prepare for their future student cohorts. Additionally, understanding the literature on "at-risk" students has a bearing on researchers. It can provide clearer criteria for selecting relevant reading material and advancing research in this field. This can facilitate synthesis and integration of findings and support systematic and meta-analysis. A comprehensive understanding of the term "at-risk" student may also inform appropriate search strategies and screening methods to better understand these student populations. In light of this, a scoping review is, therefore, about the synthesis of research that aim to map literature on a selected topic to the identification of the population of articles, the key concepts, and the context of the knowledge domain thereof [1]. It explicates the evident gaps in the knowledge domain while pinpointing common characteristics of the available evidence.

This study defines an "at-risk" student from an academic perspective as one who is likely to dropout, stop-out, burn out, or fail to complete their higher education program [3]. A dropout is someone who permanently quits without earning a degree [9]. A stop-out temporarily stops with intent to return [9]. Burning-out is a response to stress through exhaustion and low productivity [14]. Failing is when a student completes a program but does not meet the set performance standards [9]. Understanding the concept of "at-risk" students is important for Higher Education Institutions to learn from one another. It supports the motto adopted by most Higher Education Institutions during the COVID – 19 pandemic that “no one should be left behind”. Most importantly, understanding this concept will likely inspire further research towards institutional success.

The literature on "at-risk" students in higher education identifies several categories of indicators, including low pre-entry marks into Higher Education Institutions [14], poor

grade point average at entry level [17], or muffled interview scores at entry into Higher Education Institutions [18]. Some indicators connote prior experience before entry into Higher Education Institutions [17], prior acquaintance with the chosen program and career goals [19], or prior intention to dropout [21]. A lot more indicators suggest dwindling performance in tests after enrolling into Higher Education Institutions [17, 38], negative behavior during classes [21], extended exhaustion levels during lectures [19], the general extent of satisfaction with education [21], little effort exerted on tasks [21], poor study skills, as well as poor attendance [11], and poor participation in class [13, 20]. Other factors that mildly feature as indicators of "at-risk" students include lack of student support strategies [17] at institutional levels, students' demographics [17], problems related to resource allocation [33, 38], other educational barriers [13], emotional intelligence [34, 35], and learning behavior [36, 37] of the student. Institutional strategic plans [24, 25, 28] and decision-making approaches [24] are also singled out. It is alluded that Higher Education Institutions that lack proper strategic planning veiledly marginalize [13] and stigmatize [26] "at-risk" students.

Early intervention is highly regarded as the prominent way to assist "at-risk" students. Evidence supports the premise that identifying an "at-risk" student early simplifies the identification of the barriers which the student needs to overcome [2]. In fact, implementation of individualized support programs increases the probability of student success, especially when the causal factors for being at risk are correctly identified. More so, use of individualized support programs such as student counselling or peer tutoring promotes the sharing of "at-risk" students' specific risk information which can facilitate timely intervention [39] at a lower cost [40]. Thus, when attempting to understand "at-risk" students, focus should rather be on getting to know the student before attempting to solve the underlying challenges. Although direct intervention programs dominate the list of remedies for being at risk, some literature also connotes indirect interventions as tantamount, such as the need for the proper sequencing of courses and logical arrangement of the content covered in the courses that put students at risk [41]. Although indirect intervention influences the performance of students, especially in Science, Technology, Engineering, and Mathematics (STEM) degrees [41], proper intervention follows appropriate identification of the main causes of students being at risk. Many institutions are focusing on student data to improve success rates and identify "at-risk" students proactively. The goal is to reframe and broaden the understanding of "at-risk" students and address the factors that put them at risk equitably. This scoping review summarizes research on "at-risk" students, highlighting variations in research design, participants, standards, and findings. The aim is to understand the evidence and its data basis.

1.1. Objectives

Three objectives summarize this scoping review in the sequence they are presented as follows:

- The study identifies prevalent categories of “at-risk” students in the context of Higher Education Institutions. It is crucial to comprehend the concept of "at-risk" students in higher education to identify potential hurdles and ensure their success and retention through support. This helps to ensure equal educational opportunities and improve outcomes for underrepresented or disadvantaged student groups.
- The study also pinpoints the aims, analytics tools, variables, and methods insinuated when the topic on “at-risk” students is tabled. It is important to pinpoint the data analytics tools used in identifying "at-risk" students because these tools will affect the accuracy and effectiveness of the identification process. Choosing the right tools can help ensure that relevant data is collected and analyzed, leading to more accurate and actionable insights about which students may be struggling. This can inform targeted interventions to support these students and improve their outcomes.
- The study also seeks to understand what matters, where, when, and how so, when this topic on “at-risk” students is studied. Knowing this is important when discussing "at-risk" students because it allows for a more targeted and effective intervention plan while also determining the root cause of the risk factors. Such an understanding helps allocate resources more effectively while enabling the tracking of progress and success.

1.2. Research Questions

Three research questions are aligned to the given objectives as follows:

- RQ1: What is known from the existing literature about at-risk students in higher education?
- RQ2: What is the nature of the evidence relevant to the provision of interventions for at-risk students?’ and
- RQ3: What types of methodologies have been reported that seek to evaluate At-Risk students, and what approaches have been used to improve at-risk students in higher education?

The achievement of these three objectives and the response to the research questions would validate the hypotheses and theories in the literature. It will improve our understanding of 'at-risk' students while also generating new knowledge and advancing the field. This is because they identify the different types of evidence that are accessible in the literature about "at risk" students and the important traits or components that pertain to the idea of a "at risk" student, henceforth they are ideally suited for scoping reviews. Furthermore, this may create the potential to identify gaps in the literature on 'at risk' students and, therefore, present opportunities for further research aimed at promoting social justice in higher education institutions while driving data-driven institutional planning.

1.3. Overview

The rest of the article proceeded as follows; section 2 presents the methods we followed in completing this scoping review, emphasizing how the PCC (Population, Concepts, Context) framework fits into this study in guiding the selection of literature that befit the concept and context of “at-risk” students in Higher Education Institutions. The eligibility criteria, information sources, search strategy, screening procedure, and how the summaries were drawn are also described in this section 2. Section 3 presents the results which report the findings. The discussions follow in section 4, before we draw the conclusions in section 5, highlighting the main contributions of the work, and the direction for further studies.

2. Materials and Methods

2.1. The PCC framework

The thrust of this study was to categorize research on "at-risk" students in Higher Education Institutions through a scoping review. A search strategy was proposed to identify relevant studies on the topic. The study used the PCC (Population, Concept, and Context) framework, which outlines a plan for determining relevant information in a population of studies [31]. This approach helped identify the common categories of "at-risk" students in Higher Education Institutions. The search strategy defined by the PCC framework filtered relevant articles based on the Higher Education Institutions setting. The definition of "at-risk" students was demarcated to academic risk, which include students who may be at risk of dropping out, stopping out, burning out, or failing from a Higher Education Institutions perspective. A scoping review on "at-risk" students should focus on academic risks because the term "at-risk" typically refers to students who are in danger of not succeeding academically, mainly due to learning difficulties. Also, this is the only domain of “at-risk” students that Higher Education Institutions can directly mitigate. By focusing on academic risks, this scoping review provided information about the specific challenges faced by students in the classroom, as well as potential strategies and interventions that can be used to support these students’ academic success. Other forms of “at-risk” students are outside the scope of this study because Higher Education Institutions may not have the capacities, resources, and expertise to remedy the situation, such as financial risk, psychological risk, social risk, physical risk, and personal risk. This study emphasizes evidence that characterize the population, concept, and context of students under the academic risk domain.

2.2. Inclusion and exclusion criteria

The Joanna Briggs Institute (JBI) scoping review protocol [8, 31] was adopted, where those articles characterized by the keywords identified as top trending categories of

"at-risk" students (see Table 2 below) were nominated. The JBI protocol maps and summarizes the extent, range, and nature of existing research on a particular topic, to identify gaps and inform future research. Concepts such as intervention, at-risk, failing, dropout, stop-out, burn out, performance, and success were designated to define the relevant articles for this scoping review. While the key concept remained the term "at-risk" student, the context, conversely, was persistently about students in the academic risk domain, those characterized by being at risk of dropping out, stopping out, burning out, or failing in Higher Education Institutions. Thus, academic performance was an important parameter and factor throughout [27]. In this case, Higher Education Institutions refers to universities because the scope of the review is limited towards influencing institutional change and social justice at this level.

Inclusion also considered articles that were from peer-reviewed conferences and journals. Only those articles that were published in English were contemplated because of our English proficiency and the likely wider audience and distribution of English-language content. Similar studies can be conducted on articles in other languages. Articles were sampled from the EBSCOhost and ScienceDirect databases, soliciting articles that were published after the year 2010. These two databases are large enough to provide a significant sample of articles, including a diverse collection of articles and journals, and are filtrable for refined results that can be mapped to envisaged outcomes when more databases are considered. The ability to access full-text articles, organize, and cite them further intrigued us to utilize these databases. Moreover, updating these databases with fresh content and peer-reviewed sources is straightforward.

Focus was on articles published after the year 2010 to reflect more recent developments and advancements in the field, and provide updated information, data and analysis that are not available in older articles. Besides incorporating newer technologies and methodologies that could impact our research, articles less than twelve years old may address current challenges and issues that were not present before. Although this does not mean that older articles are ignored, including recent articles can help ensure that our research is up-to-date and relevant. The deep inner type of the articles reviewed was not of interest. Therefore, review articles, conceptual papers, theoretical articles, or empirical quantitative and qualitative studies were all relevant. To ensure quality, an iterative approach was adopted which allowed repeated refinement of the inclusion and exclusion criteria. Thus, articles went through several iterated screening rounds before we generated the final list of relevant literature. The research team members subjected the disputed categories to round-robin evaluations until they arrived at a consensus. Sometime, detailed manual scrutiny of the full texts of the articles were pondered on as the last resort.

2.3. Search process

We used the *litsearchr* R package [28] to facilitate quick, objective, and a reproducible search process based on text-mining and keyword co-occurrence networks

[28]. This search process minimized possible bias in the selection of relevant articles by limiting reliance on subjective and predetermined factors. Use of an automated search tool improved search recall by exploiting the identification of synonymous terms that research team members would otherwise manually miss. Also, automation of the search process took away the likely bias of researchers typically selecting keywords based on their own knowledge without specifying how the search process was administered [10]. Such bias instigates irreproducibility because it would be hard to recall the subjective procedures followed. Also, an automated tool is faster, efficient, consistent, and accurate in data retrieval. The tool has large scale search capabilities and ability to search multiple databases simultaneously, systematically, and comprehensively. This can save time and effort compared to, otherwise, manual search. The following search query, embedded in the proposed search tool, was key in mining relevant studies from the EBSCOhost and ScienceDirect databases.

$$(students \wedge at\text{-}risk \wedge (fail\ stop\text{-}out \vee burn\text{-}out \vee drop\text{-}out) \wedge (university \vee college))$$

The validity of this search query was verified with the help of an experienced librarian to improve the accuracy and relevance of the search results. Librarians have training in researching and have access to a wide range of resources and databases, which can aid in identifying the most effective search terms and strategies for a particular topic. They can refine and narrow the search, ensuring that the results are focused and relevant to the research question. Consultations with content experts in the field of student success were also considered to triangulate the meanings drawn from the searched literature, as well as to enhance rigour and reliability in the search process. Content experts were also valuable in identifying additional literature that was hard to identify through other means. The technical syntax that brought up the critical aspects of interest in the search was developed from the search query to produce the following search strategy:

"\(\("at-risk student"\∨\"at-risk"\^(dropout∨stopout∨burnout∨fail)^(college∨university\)\)"

Executing this search strategy produced a list of articles which were then subject to a targeted screening process, as outlined in the following section.

2.4. Screening of articles

The standard procedure for verifying scientific material is through manual screening. Generally, this screening can be split into several steps, including screening articles by title, abstract or physically going through the full text of the article. In our case, we automated the screening process by developing a tool using the *revtools* R package [16] that supports evidence synthesis by importing bibliographic data from the EBSCOhost and ScienceDirect databases. The *revtools* R package was used to find duplicates and remove

them (i.e., deduplicating bibliographic data). Furthermore, the tool was used for title and abstract screening. Additionally, the tool was custom-built to perform advanced screening of duplicate items that would have been missed by using topic modelling. In using the *revtools* R package, we were guided by the PRISMA-ScR framework [15]. PRISMA is an acronym for Preferred Reporting Items for Systematic reviews and Meta-Analyses [15], while ScR stands for Scoping Review. This is a tool to help authors, editors and reviewers evaluate the quality and reporting of systematic reviews to ensure completeness, transparency, unbiased opinion, and promoting best practices in the reporting of systematic reviews. In this case, the PRISMA-ScR framework facilitated the construction of a flow diagram that shows how screening was undertaken through the different stages. The flow diagram mainly reported the numbers of articles considered, included, and excluded in each step, together with the reasons for inclusion or exclusion.

2.5. Data charting process

A standardized data extraction template that followed the PRISMA-ScR format was created as part of the data charting process. Data charting in a scoping review has several benefits including visualization of data, identification of patterns, comparison of results, facilitation of data synthesis, and improved accuracy. Overall, charting simplify, clarify, and synthesize information, leading to improved understanding and decision-making. We indicated that the population of articles that met the inclusion criteria for the concept of “at-risk” student in the context of failing, dropping-out, stopping-out, or burning-out in Higher Education Institutions, together with the details of those articles in terms where the studies were conducted, when, why, with who, how, and so what, were the key results of interest. The value of extracting these details was to establish the likely knowledge gaps to explore in further research on tailored or contextualized studies on “at-risk” students, for example, in our own Higher Education Institution. The information gathered also aimed to paint a comprehensive picture of the important factors to consider when exploring the problem of "at-risk" students, including the areas where this has been a topic of concern, the timing, reasons, and methodology behind it. The primary means of presenting the likely gap maps was through the use of figures, diagrams, charts, and tables [12]. The expected outcomes were organized by region, objectives, study participants, year of study, methods used, data analytics tools employed, and the resulting findings.

2.6. Data items

The focus of this scoping review was on the twelve variables listed in Table 1. These are the key features upon which literature was synthesized and characterized. It is hoped that the abstraction of these variables from the studies that met the inclusion criteria for this scoping review will provide a clearer and more comprehensive understanding of the key factors to consider when studying the topic of "at-risk" students, including where it is most relevant, when, why, and how. This information will be valuable as a foundation

before embarking on more tailored research within the context of our own higher education institution.

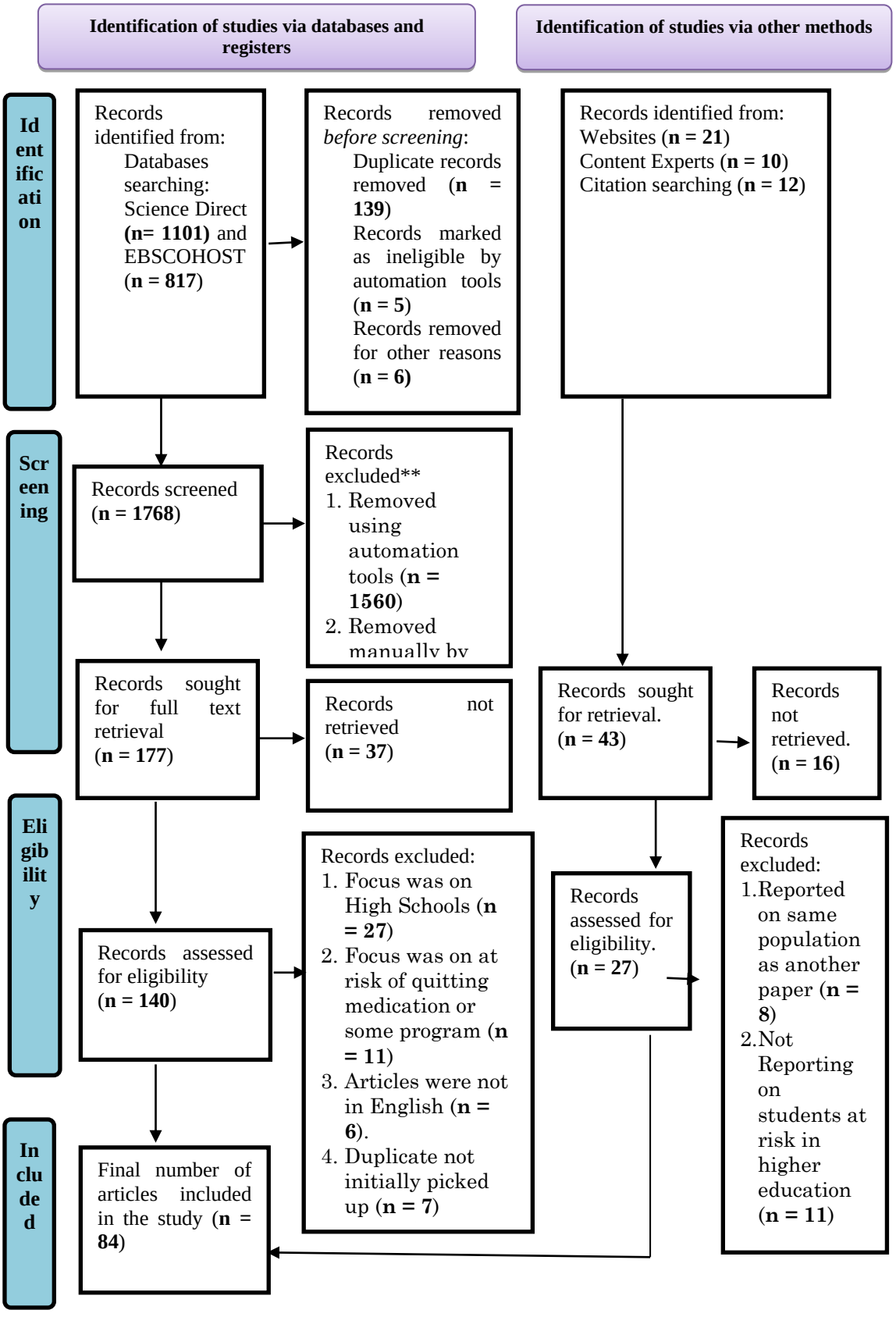
Table 1. Variables that characterized relevant literature.

Variable	Description
Author	This variable investigated the details of the authors, focusing on their names. The hope was to identify father-figures in this research area.
Country	The variable solicited the country or countries where the study was conducted or located. This would give an indication of where an understanding of "at-risk" students has been of interest.
Year of publication	This variable sought the year in which the study was done and published, thus giving an indication of how current the study was.
Educational setting	This variable was used to capture the nature of the Higher Education Institutions from which the participants were sampled. Our study focused on conventional university or blended settings.
Aims	In each study that met the inclusion criteria, we sought the key aims and objectives that the addressed. This variable would give insights of "what mattered?" and "why so?".
Study design	This variable sought to capture the type of study such as case study, survey, systematic review, qualitative, quantitative, etc. The purpose of this consideration was to respond to the "how?" aspect when attempts to understand "at-risk" students were made.
Study population	This feature was used to indicate whether the study focused on all the students in the institution or focused on a particular group of students, for example first years, or students in a particular program, or course, etc. This variable sought to respond to the "with what?" aspect of the studies.
Features	This variable extracted the characteristics of the data that was collected about the participants and how that data was useful in the study. Some examples of the features prevalently collected included students' gender, high school background, academic performance, marks, etc. This aspect also sought to respond to the "with what?" question of each study.
Data analytic tools	This variable was used to capture the approach that was embraced for data analysis. The common analytical tools ranged between statistical methods, machine learning tools, and logistic regression models. Similarly, this variable contributed to the "with what?" aspect of the study.
Interventions	This variable was used to capture the remedies that were proposed for reducing or mitigating being at risk. Precisely, this variable sought answers to the "so what?" question in each study.
Key findings	In pursuit of answers to the "so what?" question, this variable was used to record the key findings and the main conclusions drawn from each study. Hopefully, the collection of these findings will inform the meaning of the anticipated outcomes in future related studies.
Limitations	This variable captured the limitations reported from each study, if there were any. Important in this case was to give a picture of the challenges commonly encountered when studies on "at-risk" students are undertaken.

3. Results

3.1. Selection of evidence

Fig. 1. PRISMA-ScR flow diagram



The PRISMA-ScR flow diagram in Figure 1 summarizes the articles included and excluded, along with the reasons for their inclusion or exclusion. The PRISMA-ScR aims to identify articles addressing the issue of "at-risk" students in higher education institutions. The study used a *litsearchr*-based search query to extract 1918 articles from the ScienceDirect (1101 articles) and EBSCOhost (817) databases, with an additional 43 articles being added through random search (28 article), recommendations from content experts (3 articles), and citation search (12 articles). The final scoping review population consisted of a total of 1961 articles.

The first *revtools* screening that de-duplicated article using titles resulted in 139 studies being dropped. An additional five articles were eliminated through a subsequent *revtools* screening using abstracts. Six other articles were excluded as they did not align with the key concept and study context. The application of *revtools* screening with a focus on topic modeling resulted in the elimination of the largest number of articles (1560 articles). In this context, topic modeling is a statistical technique for discovering abstract topics or themes in large collections of text documents. The goal of topic modeling is to identify latent topics that are present in a corpus of text and to discover relationships between the documents and the topics. This allows for grouping of similar documents and the automatic organization of large text collections into topics.

The remaining 220 articles were reviewed manually by the members of the research team. During the process, full texts for 53 articles were unavailable, reducing the number of traceable studies to 167. These articles underwent further manual screening to verify that their content aligned with the concept and context of "at-risk" students in Higher Education Institutions . During this stage, 27 articles were rejected as their participants were outside the scope of higher education institutions. An additional 11 articles were excluded as they centered on nursing students in non-degree-granting colleges. Six more articles were discarded due to being written in a language other than English. Thirteen more articles were also not included as they dealt with different forms of risk, such as the risk of discontinuing medication or quitting other programs unrelated to education. During the full-text review process, duplicate articles were identified, and 7 articles were removed that were missed by the *revtools* application. Additionally, 8 articles were discarded as they covered the same participants as other considered articles. Finally, 11 articles were not included in the study because they had a different definition of "at-risk" students. As a result, only 84 articles were deemed relevant studies and served as the foundation for the findings, discussions, recommendations, and conclusions of the study.

It is worth mentioning that the key takeaway from this screening process is that while automated literature search (facilitated using the *litsearchr* tool) generates a large number of studies for scoping reviews, automated screening (expediated using the *revtools* application) offers speed, efficiency, objectivity, and accuracy in the screening process. The use of automated tools is therefore a valuable methodological contribution in this field of knowledge.

3.2. Characteristics of the sources of evidence

In seeking for answers to the question about pinpointing the predominant aims, analytics tools, variables, methods, and the interventions insinuated when the topic on “at-risk” students is presented, literature highlights the common factors that determine if a student is at-risk as including academic performance [17, 18], prior learning experiences [17, 19, 21], pre-entry expectations [30], personal behavior [11, 17, 20, 21, 38], and to some extent, the student's social environment [17]. It is believed that the presence of these factors in a student's profile indicates a high likelihood of poor educational outcomes [11]. These factors significantly impact academic performance. For example, supportive family, friends, and teachers, can provide motivation and encourage students to work harder and achieve their full potential. On the other hand, negative social influences, such as peer pressure, bullying, and lack of support, can have a detrimental effect on a student's academic performance, causing stress, low self-esteem, and lack of motivation. In addition, the quality of education, resources, and opportunities available in a student's community can also play a role in their academic performance. Generally, an “at-risk” student would predominantly demonstrate challenges with internalization and externalization of learning content [11]. It's worth noting that students who are considered "at-risk" will likely need intervention programs for their success [19]. These interventions should mainly focus on peer mentorship, tutoring, group studies, and improving the campus culture. The impact of non-academic factors such as the student's socio-economic background, childhood experiences, and family circumstances is not given much attention [29].

Table 2 categorizes the most significant trending topics and factors related to "at-risk" students in literature. Summaries indicate that the articles used varying terms to refer to these dominant factors. For instance, the "Grades" category encompasses factors like final exam grades, test scores, major exam marks, formative test marks, predicted grades, and prior grades. On the other hand, the “Academic” category includes factors such as academic record, motivation, support, success, performance, background, as well as academic integration. Despite this, the idea of "Grades" as a risk indicator was prevalent, with 53.8% of the articles identifying it as a factor. Other frequently mentioned concepts were "Academic" (28.6%), Gender (18.7%), GPA (13.2%), Age (12.1%), Data (11%), Course (9.9%), Race (9.9%), Study (7.7%), Support (7.7%), Time (7.7%), Semester (6.6%), Scores (5.5%), Education (5.5%), and Parent (5.5%). This finding aligns with the prominence of grades as the top trend for "at-risk" students, as reported in [17, 18].

Figure 2 shows the commonly used terms to describe "at-risk" students, with terms such as dropout, poor performance, student at-risk, failing, and success standing out. Other frequently occurring terms include academic achievement, attrition, and burnout. The literature on "stopping out" is limited. This is in line with the results from topic modeling of the dominant variables that identified the top trending factors for being at-risk. The observation of a gap in the literature on students at risk of stopping out presents an opportunity for further research and is a valuable contribution to the existing body of knowledge.

Table 2: Top trending categories of the variables

Category	**(N = 84)**	**Factors/Variables**
Grades	53.8	Final exam grades, exam scores, major test marks, marks in formative tests, predicted grades, prior grades, grades in core courses, expected course grade, grade(s), secondary school grades, quiz scores, exam scores, homework scores
Academic	28.6	Academic records, academic motivation, academic support, academic success, performance, background, academic integration
Gender	18.7	Sex
GPA	13.2	GPA
Age	12.1	Age
Data	11.0	Learner and learning data, administrative data, enrolment data, activity data, trace data, system data, learning management data
Course	9.9	Course code, course load, key courses, course observations, course non-completion, course credits, course status, Expected course grade, professor of the course, core courses
Race/ethnicity	9.9	Race, race/ethnicity
Study	7.7	Study time, study skills, study program, field of study, study results, study group, work-study
Support	7.7	Educational support, peer support, parental support, family educational support, extra educational support
Time	7.7	Interaction time with content, free time, time management, study time, travel time
Semester	6.6	End of semester survey, semester enrolled,
Scores	5.5	SAT scores, ACT scores, University Entry scores,
Education	5.5	Prior education, education values, education system, prior schooling
Parent	5.5	Parent relationships, parent occupation, parent education

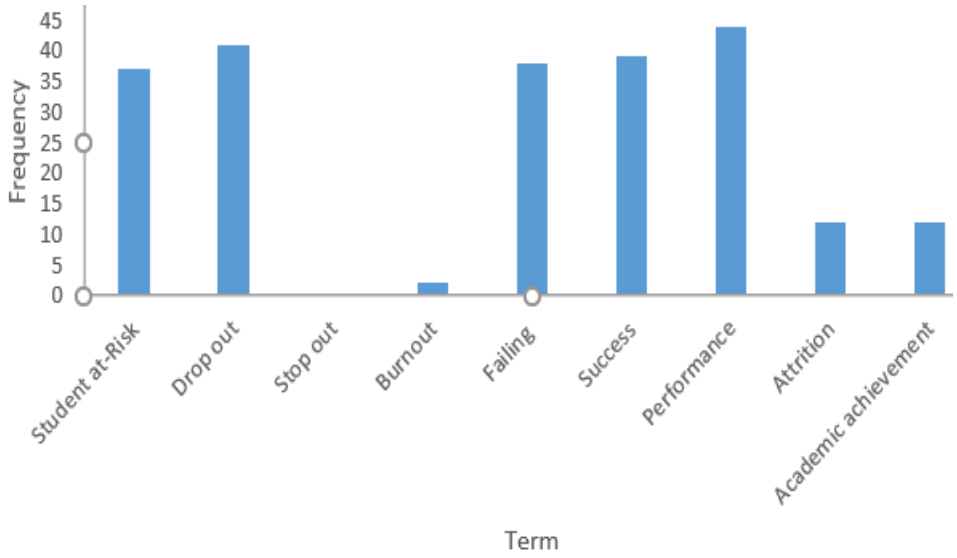


Fig 2: Popularity of the terms used to describe "at-risk" students.

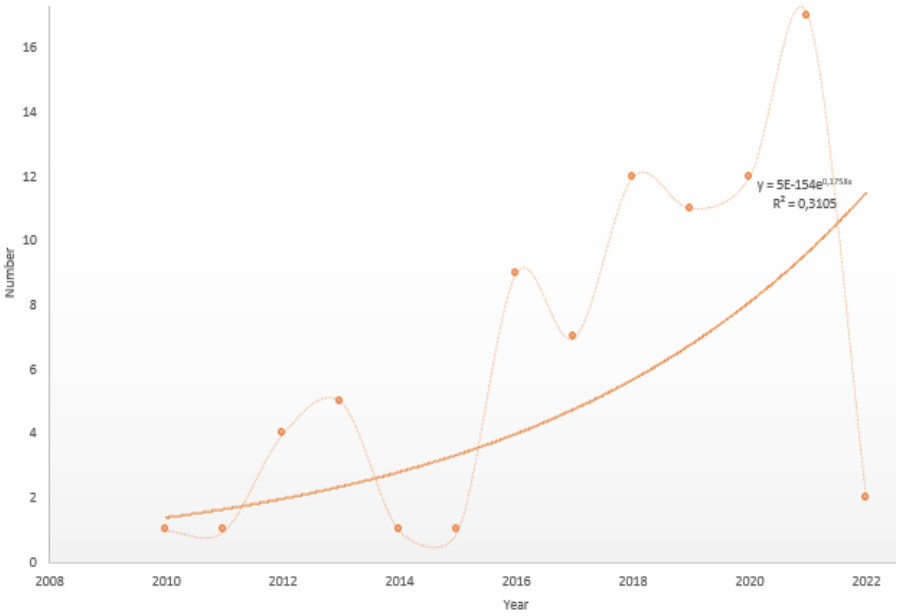


Fig 3: Articles on "at-risk" students published per year after the year 2010

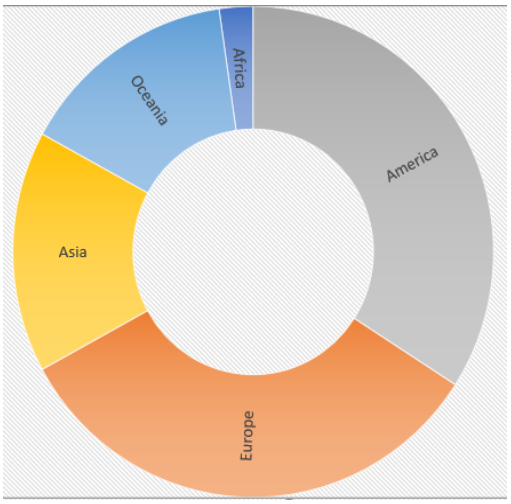


Fig. 4a: Distribution of articles by continent

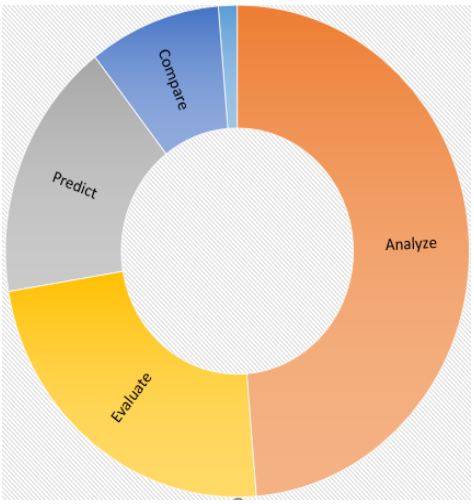


Fig. 4b: Distribution of articles by aim

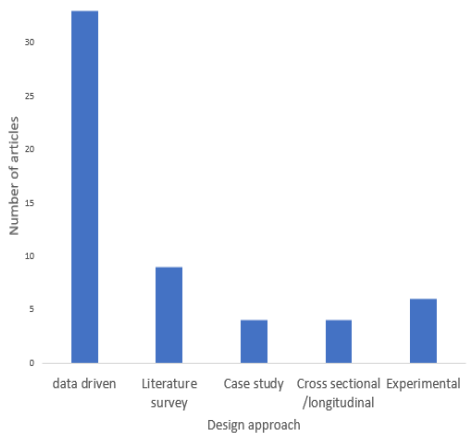


Fig. 4c: Distribution of articles by design

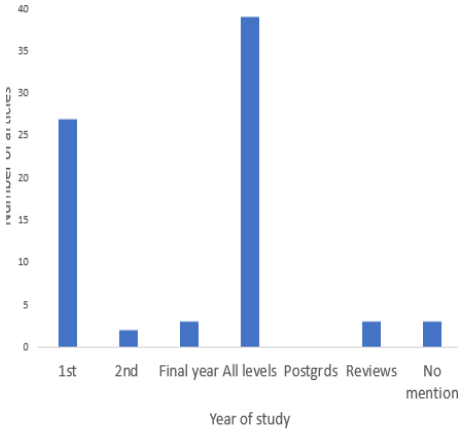


Fig. 4d: Distribution by participant level

To address the question of timing of studies (when?), Figure 3 shows the distribution of articles in this scoping review that were published between 2010 and 2022. Globally, the efforts to comprehend "at-risk" students are rapidly increasing. This increase may be due to a heightened appreciation for the positive impact that higher education institutions can have by gaining insight into their students. Most institutions that seek to understand their students often achieve good student success rates [17]. They seem to plan better and make data-informed decisions. However, the focus on this important issue is particularly noticeable in Europe, America, and Asia. Regions such as Africa are lagging (see Figure

4a), ostensibly calling for immediate intervention. This observation of the timing and locations where the topic of "at-risk" students has been of interest highlights another area worthy of exploration. Further research on "at-risk" students in regions like Africa is urgently needed.

To comprehend the reasoning behind studies on "at-risk" students, most articles focus on data analysis, evaluation techniques, or predictive models (as depicted in Figure 4b). Typically, the objective is to drive institutional progress towards undergraduate student retention and success. There are numerous other reasons for exploring "at-risk" students that can be studied in future. Although a significant portion of the literature concentrates on developing analytics, predictive, and evaluative models to identify "at-risk" students (as seen in Figure 4b), comparative studies to determine which model produces the most reliable results are scarce. This may be because the field is still in its early stages and such comparative studies may emerge in the future.

There is also very limited literature on research focusing on "at-risk" students in post-graduate studies (as shown in Figure 4d). Most studies usually target first-year students as participants, unless all students in the chosen higher education institution context are taken into account (as shown in Figure 4d). This may be due to the fact that first-year student cohorts often have the highest number of "at-risk" students. Another reason may be that the transition from high school to university is commonly seen as significant, making first-year students more in need of support than senior students. The fact that research on "at-risk" students in post-graduate studies is lacking is another gap to explore and a contribution from this study.

Data-driven methods remain popular due to the insights generated from multiple data analytics tools. Research in the form of surveys, case studies, experiments, and cross-sectional studies is also prevalent, as demonstrated in Figure 4c. However, more advanced data analytics models are preferred for making it easier to extract insights from the data collected from various information systems that institutions often use. This is likely to continue as the trend in future studies. These data analytics tools can be grouped into four broad categories (statistical methods, machine learning techniques, data mining approaches, and qualitative methods), as depicted in Figure 5. Statistical methods were the most preferred and used about 50% of the time. These methods included survival analysis, confirmatory analysis, descriptive statistics, logistic regression, multiple linear regression, cox regression, and analysis of variance. Contrary, machine learning techniques are also prevalent, accounting for roughly one third of usage. Commonly used machine learning methods include decision trees, artificial neural networks [45], naive Bayes, K-nearest neighbor, support-vector-machines (SVM), and various Ensemble techniques. Furthermore, data mining techniques and qualitative methods are also frequently used, accounting for approximately 6% and 12% respectively. Some studies utilize more than one data analytics too.

Table 3 showcases the format of the data collected for each included article. The full dataset along with all sources of evidence is available as supplementary material in a

separate spreadsheet.

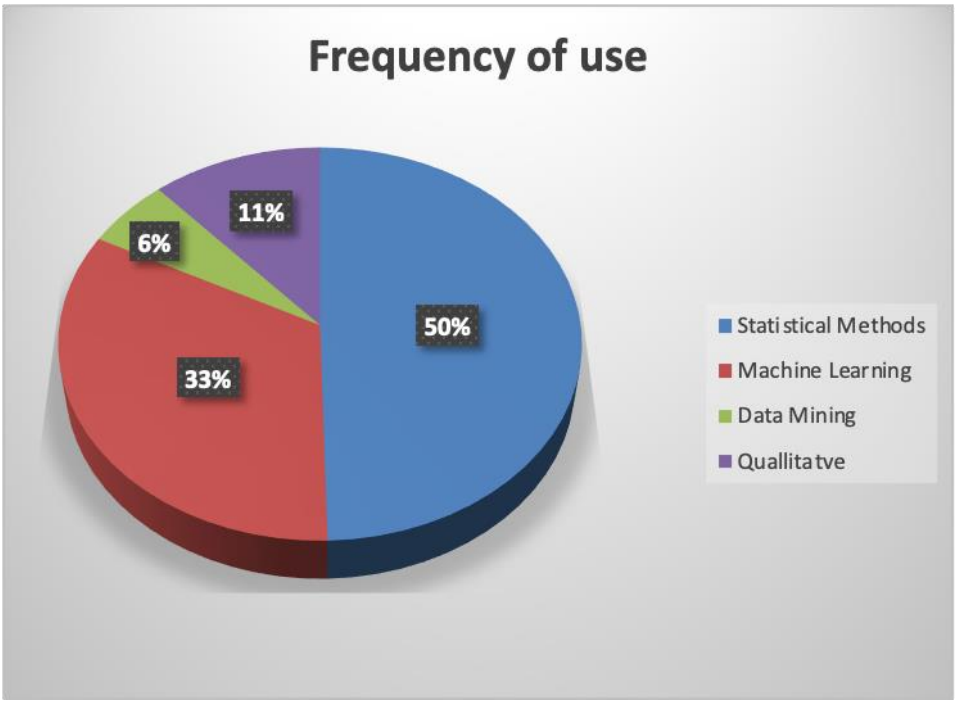


Fig. 5: Distribution of articles by the data analytics tools used.

Table 3: Format of data collected for each article.

Author	Country	Year	Setting	Aim	Design	Participants	Var	Method	Intervention
Berka, P & Marek L	Czech Republic	2021	University	Differentiate graduates from drop-outs	Data driven	First years at Czech Univ.	Family background, prior school, acadeic system, social system	Machine learning	Find hard-to-pass modules, develop a retention strategy
Paura, L and Arhipova, I	Latvia	2014	University	Analyse causes of first year dropout rates	Data driven	First years at Latvia Univ.	gender, prior school grade, source of finance	Proportional harzard model	Data from different years
Burgos, C, et.al	Spain	2018	University	Knowledge discovery to predict chance of dropout	•	Mixed group of students at Madrid Univ.	•	Logistic regression	Special tutoring action plan
• : • : • : • :	• : • : • :	• • •	• : • : • :	• : • : • :	• : • : • :	• : • : • :	• : • : • :	• : • : • :	• : • : • :

4. Discussion

RQ1: What is known from the existing literature about at-risk students in higher education?

This scoping review produced multiple results, with a recurring suggestion that low secondary school knowledge and low motivation to study can increase the risk of student dropout [42]. Students with low secondary school knowledge can enter university through various means such as remedial or catch-up programs offered by the university to help students improve their skills and knowledge in specific subjects before beginning their regular coursework. They may also enter university through alternative admission programs that allow students to demonstrate their ability to succeed in university through their life experiences or other non-academic qualifications. Some institutions offer conditional admission, where the student is admitted to the university on the condition of meeting certain academic standards within a specified period of time. In some cases, a student may just meet the academic requirements for admission to the university and enrolls in regular coursework. On the other hand, there can be a number of reasons why a student may demonstrate low motivation to study, including mere lack of interest in the subject or course material, difficulty with the coursework or feeling overwhelmed by the academic demands, personal or family issues that are affecting their ability to concentrate or focus on their studies, financial difficulties or work commitments that limit their time and energy for school, poor teaching or lack of support from instructors, negative prior experiences with education or a lack of confidence in their ability to succeed, and social or peer pressure to prioritize activities outside of school. Thus, it is important to note that every student is unique and may have a combination of factors contributing to their low motivation to study. Identifying and addressing the root cause of low motivation can help students overcome the challenges and be successful in their academic pursuits.

Being deeply invested in one's challenges can decrease the risk of a student dropping out or experiencing burnout. For example, having a positive attitude and strong commitment, combined with factors that determine cognitive abilities, demonstrate the impact of conscientiousness in reducing the likelihood of dropping out. Having self-driven motivation and effective time management skills are strong indicators of success and achievement. Similarly, there is a strong relationship between having limited high school knowledge and increased likelihood of intending to drop out, dissatisfaction with education, high levels of academic fatigue, and low expectations of graduating [4]. Low dropout rates are often associated with students who are involved in social groups [43]. Notably, funding difficulties can insinuate the impact of the student’s geographical location and ethnicity.

RQ2: What is the nature of the evidence relevant to the provision of interventions for at-risk students?’

Additionally, research suggests that interventions that provide personalized attention,

such as peer tutoring, group work, and individual counseling, are more effective. Adopting the use of early warning systems speeds up identification of students at risk. These systems strive to improve metacognitive awareness, self-awareness, and self-regulation, as well as track students' activity logs on learning management systems. That alone may simplify early prediction of “at-risk” students.

One persuasive argument is for institutions to identify courses that have a high failure rate and to assess the course material, evaluation tasks, and lecture plans to determine the perceived level of difficulty. In certain situations, it is suggested that instructors implement strategies to motivate students, such as providing prompt feedback on assignments. Furthermore, interventions that concentrate on students' psychosocial well-being and emotional intelligence are highly regarded and recommended. In this case, emotional intelligence is important for students as it helps them understand and manage their own emotions, leading to better communication, stronger relationships, and improved problem-solving abilities. Emotional intelligence can also help students cope with stress, build resilience, and make more informed decisions, which can all lead to better academic and personal outcomes.

RQ3: What types of methodologies have been reported that seek to evaluate At-Risk

students?

Higher Education Institutions can adopt machine learning models such as AutoML to create the best student performance prediction models that utilize pre-enrollment data. Using pre-enrollment data is important as it can provide valuable insights into students' background, strengths, and potential challenges, which can be used to predict their academic performance and identify areas that may need additional support. This information can help institutions tailor their services and resources to meet the individual needs of each student, which can lead to improved academic outcomes, higher student satisfaction, and reduced dropout rates. Additionally, using pre-enrollment data can also help institutions make data-driven decisions about student admission and retention, allowing them to allocate resources more effectively and efficiently.

It is also recommended to use more interpretable models that give educators insight into students' status within a course. Interpretable models can provide clear explanation of how predictions are made and why certain outcomes are generated. This makes it easier for stakeholders, such as educators and administrators, to understand the basis for the predictions and to identify areas for improvement. By having a clear understanding of the models and how they work, stakeholders can make more informed decisions and act accordingly based on the results. In addition, interpretable models can help build trust and credibility with stakeholders, who are more likely to accept and act on the results if they understand the underlying processes. Furthermore, interpretable models can also help identify any biases or inaccuracies in the data, which can be addressed and corrected to improve the overall accuracy of the predictions. Establishing a warm, encouraging, and

compassionate atmosphere in university study areas is essential for promoting a feeling of belonging.

We have noticed that while the subject of "at-risk" students is receiving significant attention, the attention to the varying interpretations of the term "at-risk" is not distributed evenly. There is a general bias towards understanding students at risk of dropping out or failing based on marks. Thus, emphasis is tilted towards interventions against dropping out or failing. Investigating the topic of students who are prone to stopping out or experiencing burnout is a clear area for further research. Likewise, many articles focused on the idea of an "at-risk" student in the context of American, European, or Asian higher education institutions. Studies on this concept from an African higher education institution perspective are tardy. It is also valuable to conduct research comparing the outcomes from various regions for the purpose of generalization.

Similarly, what gaps exist in identifying at-risk students in higher education? Although universities have various information systems that concentrate on collecting academic data, information about students' mental states is not accessible. Therefore, data about students' prior experiences, social interactions, relationships, and other extracurricular activities is necessary to deepen the understanding of "at-risk" students. In this situation, there is a gap in the existing knowledge regarding the need to examine the utilization of non-academic data to supplement the understanding of students' paths [44]. It is equally noteworthy that little is also discussed about the assessment of most suggested interventions. The effectiveness of these interventions in various higher education institution settings is not well understood, and this is another gap that merits further exploration in the future.

5. Conclusion

Examining the collection of articles that describe "at-risk" students provided insight into the objectives, participants, methods, variables, interventions, and data analytical tools that can be considered when conducting similar studies. It revealed what is important, where this has been a matter of concern, when it has been, why, and how the challenge of comprehending "at-risk" students can be addressed in various higher education institution settings. The scoping review provided insights into the gaps that are worth further exploration and the apparent methods to consider when embarking on related studies.

5.1. Contribution and value of the study

We highlight three key contributions of this scoping review to the body of knowledge:

- This research established an overview of the extensive body of literature on the

topic of "at-risk" students in higher education institutions. The literature covers important concepts and prevailing contexts. Future studies can build upon this foundation to examine the diverse student populations within specific regions and higher education institutions' settings.

- The literature was found to have several gaps, which are outlined for future research (see future works). This realization may lead to the development of new tools and innovative methods that are tailored to specific higher education institutions, particularly in the African region.
- While this scoping review was centered on comprehending "at-risk" students in higher education institutions, the results provide a foundational context from which a more comprehensive understanding of students in general can arise. Understanding students in a broad sense is an implicit responsibility for higher education institutions.

5.2. Limitations of the study

A few challenges are observed from this scoping review which may instigate further studies. These challenges are as follows:

- Comprehensive scoping reviews synthesize existing evidence on a particular topic, but the vast amount of literature covered can obscure important methodological aspects, making it challenging to define the boundaries of the studied material.
- An effective scoping process demands more time, funds, and resources that are often unpredictable at the beginning of the study. As a result, there will likely always be some overlooked elements, new perspectives, or post-hoc considerations to account for.
- Creating a search query that is thoroughly comprehensive, that reduces the number of screening rounds is challenging. The need arises for more innovative automated tools, which may be difficult to conceptualize.
- The final manual assessment of article validity for inclusion in the final round can be unrealistic in certain situations, such as when the number of articles is substantial. It's necessary to consider innovative screening methods that eliminate subjective human factors in the process.

5.3. Future direction of related studies

Various ambitious directions for future work are envisioned from this scoping review as follows:

- Studies to confirm an understanding of the "at-risk" student knowledge domain in the African context seem to be long overdue.
- Understanding an "at-risk" student as one who would likely drop out, stop out, burn out, or fail is shallow. This scoping review could be enriched by extending this context to

accommodate other use cases.

- Targeting academic data to understand students is also limited. Further research is paramount which also analyzes trace data, demographic data, and socio-economic data to better understand the broader spectrum of the student we enroll. Precisely, it is worth checking the extensibility of the concept of "at-risk" students to even consider institutional aspects.
- Focusing on understanding "at-risk" students only at undergraduate levels is also inadequate. More work is paramount towards holistically defining this concept across levels, including post-graduate levels.

Data availability

The data generated and analyzed during the study is available from the corresponding author on reasonable request.

References

Pham, M.T.; Rajic, A.; Greig, J.D.; Sargeant, J.M.; Papadopoulos, A.; McEwen, S. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Journal on research synthesis methods, John Willey and Sons. 2014. Volume 5.

Merritt, S. M. Supporting at-risk nursing students to increase their final course grade. Teaching and Learning in Nursing. 2021. Volume 16(1), pp. 74-80.

Atif, A.; Richards, D.; Liu, D.; Bilgin, A.A. Perceived benefits and barriers of a prototype early alert system to detect engagement and support 'at-risk'students: The teacher perspective. Computers & Education. 2020. Volume 156, p.103954.

Casanova, J.R.; Gomes, C.M.A.; Bernardo, A.B.; Núñez, J.C.; Almeida, L.S. Dimensionality, and reliability of a screening instrument for students at-risk of dropping out from higher education. Studies in Educational Evaluation. 2021. Vol. 68, p.100957.

Smith, B. I.; Chimedza, C.; Bührmann, J. H. Individualized help for at-risk students using model-agnostic and counterfactual explanations. Education and Information Technologies. 2021. pp. 1-20.

Cattell, L.; Bruch, J. Identifying Students at Risk Using Prior Performance versus a Machine Learning Algorithm. Regional Educational Laboratory Mid-Atlantic. 2021.

Lakkaraju, H.; Aguiar, E.; Shan, C.; Miller, D.; Bhanpuri, N.; Ghani, R.; Addison, K. L. A machine learning framework to identify students at risk of adverse academic outcomes. Proceedings of the 21st SIGKDD international conference on knowledge discovery and data mining. 2015. pp. 1909-1918.

Peters, M.; Godfrey, C.; Khalil, H.; McInerney, P.; Soares, C.; Parker, D. 2017 guidance for the conduct of JBI scoping reviews. Joana Briggs Inst Rev Man. 2017. Vol. 13. pp 141-146.

Zhang, L; Rangwala, H. Early identification of at-risk students using iterative logistic regression. In International Conference on Artificial Intelligence in Education. 2018. Springer, Cham. pp. 613-626.

Haddaway, N. R.; Woodcock, P.; Macura, B.; Collins, A. Making literature reviews more reliable through application of lessons from systematic reviews. Conservation Biology. 2015. Volume 29(6), pp 1596-1605.

McManus, E.M. Reducing School Dropout Rates Through Early Identification of Students at Risk. Emerging Perspectives: Interdisciplinary Graduate Research in Education and Psychology. 2020. Volume 4(2), pp.67-78.

Khalil, H.; Peters, M.D.J.; Tricco, A.C.; Pollock, D.; Alexander, L.; McInerney, P.; Godfrey, C.M.; Munn, Z. Conducting high quality scoping reviews - challenges and solutions. Elsevier Journal of Clinical Epidemiology. 2020. Volume 130 pp. 156 -

160.

Horton, J. Identifying At-Risk Factors That Affect College Student Success. Int. Journal of Process Education. 2015. Volume 7 (1).

Bakare, J.; Omeje, H.O.; Hisa, M.A.; Orji, C.T.; Onyechi, K.C.N.; Eseadi, C.; Nwajiuba, C.A.; Anyaegbunam, E.N. Investigation of burn out syndrome among electrical and building technology undergraduate students in Nigeria. In Medicine - clinical trials and experimental study. 2019. Volume 98:42 (e17581).

Tricco, A. C.; Lillie, E.; Zarin, W.; O'Brien, K. K.; Colquhoun, H.; Levac, D.; ... Straus, S. E. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Annals of internal medicine. (2018). Volume 169(7), pp 467-473.

Westgate, M.J. revtools: An R package to support article screening for evidence synthesis. Research synthesis methods, 2019. Volume 10(4), pp 606-614.

Russell, J.; Smith, A.; Larsen, R. Elements of Success: Supporting at-risk student resilience through learning analytics. Elsevier journal of Computers and Education. 2020. Vol (152).

Larose, S.; Tarabulsy, G.M. Academically at-risk students. Book: Handbook of Youth Mentoring. 2nd Edition, Sage Publishers. 2014.

Joana, R.C; Gomes, C.M.A; Bernardo, A.B.; Núnez, J.C.; Almeida, L.S. Dimensionality and reliability of a screening instrument for students at-risk of dropping out from Higher Education. Elsevier journal of studies in educational evaluation. 2021. Vol 68.

Zeineddine, H.; Braendle, U.C.; Farah, A. Enhancing prediction of student success: Automated machine learning approach. Comput. Electr. Eng. 2021. Volume 89.

Owen, H.T.L.; Huang, A.Y.Q.; Yang, S.J.H. Impact of teachers' grading policy on the identification of at-risk students in learning analytics. Elsevier journal of Computers and Education. 2021. Volume 163.

Mashiloane, M; Mchunu, M. Mining for Marks: A Comparison of Classification Algorithms when Predicting Academic Performance to Identify students at risk. LNAI. 2013. Volume 8284, pp. 541–552. Springer International Publishing.

Al-Shabandar, R.; Hussain, A.J.; Liatsis, P.; Keight, R. Detecting At-Risk Students with Early Interventions Using Machine Learning Techniques. 2018.

Newton, P.; Burgess, D. The Best Available Evidence: Decision Making for Educational Improvement. Springer. 2016..

Garard, D. Defining the At-Risk Student: Conceptual and Theoretical Considerations. Annual Meeting of the Speech Communication Association (81st, San Antonio, TX, November 18-21). 1995.

Lim, C.P.; Wang, P.; Graham, C. Driving, sustaining and scaling up blended learning practices in higher education institutions: a proposed framework. Innovation and

Educ. 2019

Wilton, M.; Gonzalez-Niño, E.; McPartlan, P.; Terner, Z.; Christoffersen, R.E.; Rothman, J.H. Improving Academic Performance, Belonging, and Retention through Increasing Structure of an Introductory Biology Course. CBE - Life Sciences Education. 2019. (18) (4).

Grames, E.M.; Stillman, A.N.; Tingley, M.W.; Elphick, C.S. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. Methods in Ecology and Evolution. British Ecological Society. (2019). Vol 10 (10).

Bunn, M.; Bennett, A.; Burke, P. J. In the anytime: Flexible time structures, student experience and temporal equity in higher education. Time & Society. 2014. Volume 28(5).

Berka, P.; Marek, L. Bachelor's degree student dropouts: Who tend to stay and who tend to leave? Studies in Educational Evaluation. 2021. Volume 70.

Peters, M.; Godfrey, C.; Khalil, H.; McInerney, P.; Soares, C.; Parker, D. 2015 Methodology for JBI Scoping reviews. The Joana Briggs Institute Reviewers' Manual 2015 Edition, The Joanna Briggs Institute. 2015.

Snilstveit, B.; Vojtkova, M.; Bhavsar, A.; Gaarder, M. Evidence Gap Maps: A Tool for Promoting Evidence-Informed Policy and Prioritizing Future Research. Policy Research Working paper 6725, The World Bank, Independent Evaluation Group. 2013.

Jayaprakash, S.M.; Moody, E.W.; Lauría, E.J.M.; Regan, J.R.; Baron, J.D. Early Alert of Academically At- Risk Students: An Open-Source Analytics Initiative. Journal of Learning Analytics. 2014.

Shafait, Z.; Khan, M.A.; Sahibzada, U.F.; Dacko-Pikiewicz, Z., Popp, J. An assessment of students' emotional intelligence, learning outcomes, and academic efficacy: A correlational study in higher education. PLoS ONE. 2021. Volume 16(8).

Chong, A.M.; Lee, P.G.; Roslan, S.; Baba, M. Emotional Intelligence, and at-risk Students. Sage Open. 2015. pp 1-8.

Jelena, J.; Mohammed, S.; Srecko, J.; Dragan, G. Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. Computers & Education. 2021. Volume 172.

Sarra, A.; Fontanella, L.; Di Zio, S. Identifying Students at Risk of Academic Failure Within the Educational Data Mining Framework. Soc Indic Res. 2019. Volume 146, pp 41–60.

Delmas, P. M.; Childs, T. N. Increasing faculty engagement in the early alert process. Innovations in Education and Teaching International. 2021. Volume 58:3, pp 283-293.

Brannagan, K.B.; Dellinger, A.; Thomas, J.; Mitchell, D.; Lewis-Trabeaux, S.; Dupre, S. Impact of peer teaching on nursing students: perceptions of learning environment, self-efficacy, and knowledge. Nurse Educ Today. 2013. Volume 33(11): pp 1440-7.

Kim, S. C.; Jillapali, R.; Boyd, S. Impacts of peer tutoring on academic performance of first-year baccalaureate nursing students: A quasi-experimental study. Nurse Education Today. 2021. Volume 96, pp 104658.

Zhuhadar, L; Daday, L.; Marklin, S.; Kessler, B.; Helbig, T. Using survival analysis to discovering pathways to success in mathematics. Computers in Human Behavior. 2019. Volume 92, pp 487-495.

Paura, L.; Arhipova, I. Cause analysis of students' dropout rate in higher education study program. Procedia-Social and Behavioral Sciences. 2014. Volume 109, pp 1282-1286.

Masserini, L.; Bini, M. Does joining social media groups help to reduce students' dropout within the first university year? Socio-Economic Planning Sciences. 2021. Vol. 73.

Van der Zanden, P.J.C.; Denessen, E.; Cillessen, A.H.N.; Meijer, P.C. Domains and predictors of first-year student success: A systematic review. Educational Research Review. 2018. Vol 23. pp 57-77.

Wang X., Guo, B., Shen, Y. Predicting the At-Risk Online Students Based on the Click Data Distribution Characteristics. Scientific Programming. 2022.