# Survival Analysis in Education: A Systematic Literature Review, Bibliometric Insights, and Methodological Comparison

## Abstract

Background:

Dropout and student retention are critical issues in education, with significant individual and societal implications. Survival analysis provides a powerful statistical framework for examining time-to-event phenomena such as student attrition, yet its application in educational research has not been systematically reviewed.

This study conducts a bibliometric and methodological review of the use of survival analysis in educational contexts, with a focus on identifying prevailing methods, their implementation success rates, and research trends across countries and institutions.

Methods:

We conducted a systematic literature search in the Web of Science database using keywords related to survival analysis and education. In total, 297 peer-reviewed journal articles published between 1991 and 2025 were analyzed. Abstracts were processed using Python to identify mentions of survival methods (e.g., Cox regression, Kaplan-Meier, competing risks) and indicators of model performance (e.g., fit, significance, validation). A bibliometric analysis was also performed to examine publication trends, collaboration networks, and geographical distribution.

Results:

The most frequently used methods were event history analysis (n=51), Cox proportional hazards models (n=36), and Kaplan-Meier estimators (n=27). Success in application, as indicated by mentions of model performance or validation, varied by method. Kaplan-Meier methods exhibited the highest success ratio (59%), while newer techniques like multilevel and dyadic survival models were underutilized. The majority of studies originated from North America and Western Europe, while representation from developing regions remains limited. Collaboration networks revealed strong ties among researchers in the United States, United Kingdom, and Australia.

Conclusion:

Survival analysis is increasingly used in educational research, particularly for modeling dropout and retention. However, the field relies heavily on conventional methods, with limited adoption of advanced techniques. Future research should prioritize innovation in methodology and broader global inclusion.

## Keywords

Survival Analysis, Education Research, Dropout Prediction, Student Retention, Time-to-Event Analysis, Systematic Literature Review, Bibliometric Analysis, Methodological Comparison

## Introduction

Educational attainment plays a crucial role in personal development, economic mobility, and societal progress (Padgett et al., 2025). However, dropout remains a significant issue across various educational levels, impacting both individual students and educational institutions. Understanding the dropout rates is essential for developing effective policies and interventions (Freeman & Simonsen, 2015). A robust methodological approach to studying time-dependent educational outcomes is survival analysis, a statistical framework designed to model time-to-event data. This study aims to explore the application of survival analysis in education by conducting a systematic literature review, performing a bibliometric analysis, and comparing methodological approaches within the field.

### Dropout in education

Dropout rates in education vary across different regions, demographics, and educational levels, influenced by factors such as personal and economic issues, academic performance and satisfaction, social and institutional support, psychological determinants and demographic variables (Addison & Williams, 2023; Pusztai et al., 2022; Wild & Schulze Heuling, 2020). High dropout rates can lead to long-term negative consequences, including individual reduced impact like employment opportunities (Kiprianos & Mpourgos, 2025) and societal impacts (Ressa & Andrews, 2022). Policymakers and educators have long sought ways to mitigate dropout rates through targeted interventions. However, accurately predicting and understanding dropout behaviour requires robust analytical methods that account for time-dependent risk factors and censoring, making survival analysis an ideal tool for such investigations (Wild & Schulze Heuling, 2020).

### Survival analysis methods

Survival analysis encompasses a suite of statistical techniques designed to model the time until the occurrence of a specific event, such as student dropout or graduation. In the educational context, these methods are particularly valuable due to their ability to account for censoring—cases where the event of interest has not yet occurred—and to incorporate time-varying covariates that influence risk over time (Wild & Schulze Heuling, 2020).

The most commonly used techniques include, but are not limited to:

* Kaplan-Meier Estimator (Lee, 2023): A non-parametric method used to estimate survival probabilities over time without assuming an underlying distribution.
* Cox Proportional Hazards Model (Lee, 2023): A semi-parametric model that assesses the impact of covariates on the hazard function, widely used for its flexibility and interpretability.
* Discrete-Time Survival Models (Schmid & Berger, 2021): Suitable for data collected in regular intervals (e.g., academic years), allowing for logistic regression frameworks.
* Competing Risks Models (Schmid & Berger, 2021): Applied when individuals are exposed to multiple potential events, enabling the analysis of cause-specific outcomes.

These methods enable researchers to capture the dynamic and multifactorial nature of educational trajectories, offering nuanced insights into factors that influence student persistence and attrition over time.

### Existing literature

Despite the long-standing application of survival analysis in fields like medicine and engineering, its systematic use in educational research remains relatively limited. One of the most influential and methodologically rigorous contributions is the systematic literature review by (Willett & Singer, 1991) titled "From Whether to When: New Methods for Studying Student Dropout and Teacher Attrition." This foundational work introduced educational researchers to the core principles of survival analysis, emphasizing its relevance for modelling time-to-event educational outcomes. The authors showcased a range of survival methods—such as Kaplan-Meier estimators and discrete-time hazard models—and applied them to teacher attrition and student dropout, arguing for the necessity of analysing when events occur, not just if they occur. Their review highlighted the challenges of censoring in educational datasets and offered a conceptual and methodological framework that continues to inform survival-based educational studies today​.

However, as survival analysis techniques have advanced—particularly through the integration of machine learning, expanded covariate modelling, and broader educational contexts—there is a need to revisit this body of literature through the lens of newer empirical contributions, even if they are not categorized as review articles. For that, we selected 10 of the most relevant 2025 articles—based on citation impact, methodological clarity, and substantive relevance—from an already filtered pool described in the methods section; the full list can be accessed in the Appendix A: Most relevant 2025 studies found. Recent applications of survival analysis in educational research demonstrate its utility across policy, equity, and workforce studies, though methodological and contextual gaps persist. For instance, Li & Hu (2025) employed dyadic survival analysis to model performance-based funding discontinuation in U.S. higher education, revealing political imitation as a key driver, while Harris-Walls & Curran (2025) used event history analysis to show partisan divides in K-12 science standards adoption. Studies like Shiferaw et al. (2025) and Kwon et al. (2025) advanced attrition research—applying discrete-time hazard and multilevel survival models to link professional development programs to teacher retention (Shiferaw et al., 2025) and classroom factors to Head Start turnover (Kwon et al., 2025). Corradi et al. (2025) highlighted equity concerns, finding affirmative action-admitted students in Chile faced higher dropout risks independent of academic performance. While these works reaffirm survival analysis’s value for timing-sensitive questions, they predominantly use conventional methods (e.g., Cox models) and focus on Western contexts, underscoring needs for advanced techniques (e.g., competing risks) and broader geographic representation. Accuracy and validation are rarely discussed in these studies; while Kwon et al. (2025) and Corradi et al. (2025) implicitly validate their models through covariate significance, they lack robustness checks such as comparisons with machine learning methods. Notably, Shiferaw et al. (2025) operationalize core aspects of Willett & Singer’s framework in evaluating educational programs, bridging classic theory with modern application. Still, the absence of more advanced survival techniques—such as competing risks models or time-varying covariates—in these 2025 studies points to continued opportunities for methodological development within the field.

### Research Questions

This study seeks to address the following research questions:

1. What are the key bibliometric trends in the literature on survival analysis in education, including publication output, influential authors and institutions, and prevailing research themes as identified?
2. What is the geographical distribution of research on survival analysis in education, and what patterns of international collaboration are evident within the field?
3. What are the predominant survival analysis methods used in education research?
4. With what success has survival analysis been applied to model student retention and dropout rates across various educational contexts and levels? What are the key trends and contributions in the literature, as identified through bibliometric analysis?

## Methods

### Literature search

To assess the state of research on survival analysis in education, a systematic search was conducted using the Web of Science (WoS) database. The final search was performed in April 2025 using the following query:

("survival analysis" OR "event history analysis" OR "hazard model" OR "time-to-event analysis" OR "Kaplan-Meier" OR "Cox regression" OR "discrete-time survival" OR "competing risks") AND ("education" OR "educational attainment" OR "student retention" OR "dropout" OR "graduation" OR "academic success" OR "school completion" OR "college completion" OR "student success").

The results were refined by limiting the research area to “Education” and “Educational research”, which yielded 297 results.

Article selection was conducted by a single reviewer. To enhance rigor, the screening protocol was pilot tested on a subset of 30 articles, ensuring consistent application of eligibility rules. While inter-rater reliability was not assessed formally, efforts were made to minimize bias through protocol transparency and documentation.

To ensure the comprehensiveness and robustness of our bibliometric analysis, we extended our search beyond WoS and conducted an additional query in the Taylor & Francis database. The query used a comparable structure, targeting both methodological and educational terms within the subject area "Education":

[[All: "survival analysis"] OR [All: "event history analysis"] OR [All: "hazard model"] OR [All: "time-to-event analysis"] OR [All: "kaplan-meier"] OR [All: "cox regression"] OR [All: "discrete-time survival"] OR [All: "competing risks"]] AND [[All: "education"] OR [All: "educational attainment"] OR [All: "student retention"] OR [All: "dropout"] OR [All: "graduation"] OR [All: "academic success"] OR [All: "school completion"] OR [All: "college completion"] OR [All: "student success"]] AND [All Subjects: Education]

This query yielded 367 results, reflecting a slightly broader pool of publications compared to WoS. To assess the overlap and uniqueness of these results, we employed a fuzzy matching algorithm to compare the retrieved articles against the WoS dataset. This technique allows for minor variations (e.g., one-character differences, formatting inconsistencies, or slight modifications in metadata) while still identifying potentially identical studies across databases. The results of the fuzzy matching analysis revealed that: 44 articles (12%) were a 100% match between the two databases. 48 additional articles (13%) exhibited a 90% or higher similarity score, indicating potential duplication with slight variations in metadata. The remaining 275 articles (75%) did not meet the matching threshold, suggesting a relatively low overlap between Taylor & Francis and WoS datasets. These findings highlight that while Taylor & Francis includes a number of relevant articles, its overlap with WoS is limited, emphasizing the importance of multi-database searches in conducting a comprehensive systematic review. The relatively low alignment percentage also suggests a degree of publication exclusivity between these databases, which may be influenced by indexing policies and journal coverage differences.

### Data analysis

The final set of 297 articles was analysed using Python, incorporating custom-built scripts for bibliometric visualization. While existing libraries such as pyBibX offer robust tools for bibliometric and scientometric analysis, we developed a tailored solution to better align with the specific needs of our study. Selected functions from pyBibX were integrated into our pipeline to support core tasks such as citation mapping, co-authorship analysis, and keyword clustering.

## Results

The final dataset consisted of 297 documents, with no duplicates, covering a timespan from 1991 to 2025. The majority of the records were peer-reviewed journal articles (256), followed by articles in press (15), conference proceedings (22), book chapters (2), and review papers (2). The dataset spanned 50 countries and 302 institutions, with research published in 110 different sources. The dominant language was English (290 documents), with a few studies published in German (2), Spanish (3), Portuguese (1), and Dutch (1).

In terms of bibliometric completeness, most entries contained abstracts (98.99%), author affiliations (99.66%), and DOIs (93.60%), while 78.45% included author keywords and 88.55% had Keywords Plus. Collaboration was notable, with 768 authors contributing to 297 documents, leading to an average collaboration index of 2.94. Of these, 62 were single authored while 235 were multi-authored.

Regarding impact, the dataset accumulated 7,015 total citations, translating to an average of 23.62 citations per document. The highest recorded h-index was 7, indicating a moderate level of influence within the field. Additionally, citation distribution revealed an average of 9.13 citations per author, 23.23 per institution, and 57.3 per source.

To assess whether our search strategy effectively captured the key themes of survival analysis in education, we conducted a word cloud analysis of the extracted dataset. The word cloud in Figure 1 visualizes the most frequently occurring terms in the titles, abstracts, and keywords of the selected studies. Larger words indicate higher frequency, reflecting dominant topics within the dataset. The results confirm that our search query successfully encompassed the core concepts of survival analysis and education. The most prominent terms include "student," "study," "school," "time," "survival," "analysis," "teacher," "dropout," "retention," and "academic"—all of which align with our research focus. The presence of terms related to statistical methods (e.g., "model," "risk," "data," "event," "rate") further validates that the dataset is methodologically relevant. Additionally, words like "higher education," "university," and "college" indicate that the dataset captures various levels of the education system, ensuring broad applicability.

A close-up of words

AI-generated content may be incorrect.

Figure : Word cloud analysis.

To address the predominant survival analysis methods used in educational research, we conducted a text mining analysis on the abstracts. We searched for specific keywords and phrases associated with known survival analysis techniques (e.g., "Cox Proportional Hazards", "Kaplan-Meier", "Event History Analysis", etc.). Each abstract was scanned using regular expressions to detect mentions of each method, while simultaneously identifying references to success or validation through terms like accuracy, valid, significant, performance, and results supported, summarized in Table 1. Regular expressions code-words used can be found in Appendix B: Regular Expressions Used for Method and Success Keyword Detection.

Table : Frequency of survival analysis methods and associated success mentions.

|  |  |  |  |
| --- | --- | --- | --- |
| Method | Article Count | Success Mentions | Success Ratio |
| Event History Analysis | 51 | 17 | 0.33 |
| Cox Proportional Hazards | 36 | 18 | 0.50 |
| Hazard Model | 27 | 13 | 0.48 |
| Kaplan-Meier | 27 | 16 | 0.59 |
| Discrete-Time Survival | 20 | 9 | 0.45 |
| Competing Risks | 16 | 6 | 0.38 |
| Parametric Models | 4 | 4 | 1.00 |
| Machine Learning Survival | 3 | 1 | 0.33 |
| Time-varying Covariates | 2 | 1 | 0.50 |
| Dyadic Survival | 1 | 0 | 0.00 |
| Multilevel Survival | 1 | 0 | 0.00 |

The results reveal that Event History Analysis, a broad framework encompassing multiple survival analysis techniques, was the most frequently referenced category, followed by specific methods such as Cox Proportional Hazards and Kaplan-Meier estimators. Notably, Kaplan-Meier had the highest ratio of success mentions among the commonly used methods (59.3%), indicating favourable outcomes or interpretations in reported applications. Parametric models, though used in only four studies, showed a 100% success mention rate, albeit with limited generalizability due to small sample size. On the other hand, more advanced or nuanced techniques—such as machine learning-based survival models, time-varying covariates, multilevel, and dyadic survival models—appear underutilized in the current literature. Their low frequency, coupled with low or no reported validation mentions, indicates potential areas for methodological development in future research.

To analyse the geographical distribution of research on survival analysis in education, we mapped the number of documents produced by each country, considering the affiliations of all contributing authors, as shown in Figure 2 . Countries shaded in darker colours indicate higher research output, while lighter colours represent lower contributions. This analysis highlights the uneven distribution of research output, suggesting that survival analysis methods in education are predominantly explored in North America and select European countries, while developing regions remain underrepresented.

A map of the world

AI-generated content may be incorrect.

Figure : Figure 2. Geographical distribution of publications on survival analysis in education based on author affiliations, where the count reflects the number of documents associated with each country considering all contributing authors; darker shades represent higher research output.

The collaboration network presented in Figure 3 illustrates the global research partnerships in survival analysis applied to education.

A map of the world

AI-generated content may be incorrect.

Figure : International collaboration network.

European countries exhibit dense collaboration networks, particularly among the United Kingdom, Germany, France, Portugal, and the Netherlands. North America (United States and Canada) plays a central role, collaborating extensively with European and Australian researchers. Australia emerges as another key player, engaging in significant partnerships with both Asian and European countries. South America and Africa show some engagement, but their collaboration networks are relatively sparse, suggesting limited international research connections. India and Pakistan stand out as key contributors from Asia, actively collaborating with other regions.

The strong interconnectedness of researchers from Western countries suggests a well-established academic exchange, while developing regions show fewer international collaborations. These findings highlight the potential for expanding global partnerships, particularly by fostering research collaborations with underrepresented regions to diversify methodological approaches and broaden the scope of survival analysis in education.

## Discussion

In this article we provided a comprehensive overview of how survival analysis is applied in educational research, particularly in the study of student retention, dropout, and progression. The results highlight several important methodological and geographical trends.

First, the findings show that a core set of classical survival analysis methods—notably Event History Analysis, Cox Proportional Hazards, and Kaplan-Meier estimators—dominates the field, highlighting their accessibility and interpretability for educational researchers (Beg, 2024; Langbein et al., 2024).

While the prevalence of these traditional techniques may reflect their robustness and familiarity, the relative scarcity of machine learning survival models, multilevel, and dyadic survival models points to underexplored methodological frontiers. This is particularly relevant given the increasing complexity of educational data and the call for personalized, time-sensitive interventions (Baker & Inventado, 2014; Jarke & Breiter, 2019). Our findings suggest a gap between methodological innovation and adoption, indicating that the field may benefit from more training, dissemination, and validation of advanced survival modelling techniques.

The presence of success-related keywords in many abstracts provides an additional layer of insight. While a high success ratio in some methods (e.g., parametric models) may reflect their appropriateness for specific datasets or contexts, such interpretations should be treated cautiously due to small sample sizes and limitations in keyword-based identification. Still, the relative frequency of success mentions may indirectly reflect researcher satisfaction, model fit, or acceptance by peer reviewers, and thus deserves further exploration.

Geographically, the dominance of research output from North America and Western Europe aligns with previous patterns in education and social sciences research. However, the emerging contributions from countries such as China, South Korea, and Brazil suggest a globalizing interest in quantitative approaches to educational equity and progression. This expanding diversity may also introduce new applications and interpretations of survival models across different educational systems. Preferences for statistical methods often vary by country, shaped by national research traditions and contextual factors, as demonstrated in non-education fields—for example, in bankruptcy prediction research by (Kovacova et al., 2019).

Overall, this study offers a structured overview and guidance for both new and experienced researchers seeking to navigate the methodological landscape of survival analysis in education. By synthesizing current trends, this review helps identify well-trodden paths, as well as unexplored terrain that could yield valuable insights in the years ahead.

## Limitations and future directions

This study is limited by its reliance on abstracts, which may omit detailed methodological descriptions or success indicators. Keyword matching may have missed variations in phrasing, and the analysis was restricted to English-language, indexed publications.

Future research should explore full-text analysis to capture richer methodological context, apply more advanced text mining techniques for classification, and examine the quality and rigor of survival method implementation across diverse educational settings.

## Data availability

All data and statistical code used in this study are publicly available at https://github.com/author

## References

Addison, L., & Williams, D. (2023). Predicting student retention in higher education institutions (HEIs). *Higher Education, Skills and Work-Based Learning*, *13*(5), 865–885. https://doi.org/10.1108/HESWBL-12-2022-0257

Baker, R. S., & Inventado, P. S. (2014). Educational Data Mining and Learning Analytics. In J. A. Larusson & B. White (Eds.), *Learning Analytics* (pp. 61–75). Springer New York. https://doi.org/10.1007/978-1-4614-3305-7\_4

Beg, Dr. S. W. (2024). METHODOLOGIES IN PREDICTIVE ANALYSIS OF MEDICAL EDUCATION: A COMPREHENSIVE REVIEW. *Era’s Journal of Medical Research*, *11*(1), 56–59. https://doi.org/10.24041/ejmr2024.9

Corradi, B., Espinosa, D., Rodríguez, C., & Espinoza, O. (2025). Is Admission Enough? University Persistence of Students Admitted Through Affirmative Action Policies in Chile. *Higher Education Policy*. https://doi.org/10.1057/s41307-024-00393-w

Freeman, J., & Simonsen, B. (2015). Examining the Impact of Policy and Practice Interventions on High School Dropout and School Completion Rates: A Systematic Review of the Literature. *Review of Educational Research*, *85*(2), 205–248. https://doi.org/10.3102/0034654314554431

Harris-Walls, K., & Curran, F. C. (2025). An Event History Analysis of the Policy Adoption of the Next Generation Science Standards. *Educational Policy*, 08959048251339120. https://doi.org/10.1177/08959048251339120

Jarke, J., & Breiter, A. (2019). Editorial: The datafication of education. *Learning, Media and Technology*, *44*(1), 1–6. https://doi.org/10.1080/17439884.2019.1573833

Kiprianos, P., & Mpourgos, I. (2025). School failure and life trajectories. *Journal of Adult and Continuing Education*, 14779714251314681. https://doi.org/10.1177/14779714251314681

Kovacova, M., Kliestik, T., Valaskova, K., Durana, P., & Juhaszova, Z. (2019). Systematic review of variables applied in bankruptcy prediction models of Visegrad group countries. *Oeconomia Copernicana*, *10*(4), 743–772. https://doi.org/10.24136/oc.2019.034

Kwon, K.-A., Jang, W., Ford, T. G., Horm, D., Yazejian, N., & Bryant, D. (2025). A longitudinal study of Head Start teacher turnover trends and factors. *Teaching and Teacher Education*, *156*, 104916. https://doi.org/10.1016/j.tate.2024.104916

Langbein, S. H., Krzyziński, M., Spytek, M., Baniecki, H., Biecek, P., & Wright, M. N. (2024). *Interpretable Machine Learning for Survival Analysis* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2403.10250

Lee, S. W. (2023). Kaplan-Meier and Cox proportional hazards regression in survival analysis: Statistical standard and guideline of Life Cycle Committee. *Life Cycle*, *3*, e8. https://doi.org/10.54724/lc.2023.e8

Li, A. Y., & Hu, X. (2025). Goodbye Performance-Based Funding: Policy Abandonment of Performance Funding for Higher Education in the United States. *Research in Higher Education*, *66*(1), 8. https://doi.org/10.1007/s11162-024-09828-0

Padgett, R. N., Le Pertel, N., Hinton, C., Kettler, T., Ho, A. D., Chen, Y., Bradshaw, M., Johnson, B. R., & VanderWeele, T. J. (2025). A cross-national analysis of sociodemographic variation in educational attainment. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-87406-9

Pusztai, G., Fényes, H., & Kovács, K. (2022). Factors Influencing the Chance of Dropout or Being at Risk of Dropout in Higher Education. *Education Sciences*, *12*(11), 804. https://doi.org/10.3390/educsci12110804

Ressa, T., & Andrews, A. (2022). HIGH SCHOOL DROPOUT DILEMMA IN AMERICA AND THE IMPORTANCE OF REFORMATION OF EDUCATION SYSTEMS TO EMPOWER ALL STUDENTS. *International Journal of Modern Education Studies*, *6*(2). https://doi.org/10.51383/ijonmes.2022.234

Schmid, M., & Berger, M. (2021). Competing risks analysis for discrete time‐to‐event data. *WIREs Computational Statistics*, *13*(5), e1529. https://doi.org/10.1002/wics.1529

Shiferaw, M., O’Hagan, K. G., & Weinstein, M. (2025). Staying Put: Positive Spillovers on Teacher Retention From a Middle School Science Initiative. *Education and Urban Society*, *57*(5), 462–486. https://doi.org/10.1177/00131245251315180

Wild, S., & Schulze Heuling, L. (2020). Student dropout and retention: An event history analysis among students in cooperative higher education. *International Journal of Educational Research*, *104*, 101687. https://doi.org/10.1016/j.ijer.2020.101687

Willett, J. B., & Singer, J. D. (1991). From Whether to When: New Methods for Studying Student Dropout and Teacher Attrition. *Review of Educational Research*, *61*(4), 407–450. https://doi.org/10.3102/00346543061004407

## Supplementary files

### Appendix A: Most relevant 2025 studies found

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Article Title | Authors | Method | Level | Key findings | Relevance |
| Goodbye Performance-Based Funding: Policy Abandonment of Performance Funding for Higher Education in the United States | Li, A. Y.; Hu, X. D. | Dyadic survival analysis | Higher education policy | PBF discontinuation driven by political imitation, not geography or economics | Expands policy diffusion theory to discontinuation |
| An Event History Analysis of the Policy Adoption of the Next Generation Science Standards | Harris-Walls, K.; Curran, F. C. | Event history analysis | K-12 policy | Republican-majority legislatures prefer partial NGSS adoption over full adoption | Highlights political barriers to standardized reforms. |
| Overlapping Demographic Backgrounds and Higher Educational Attainment: Measurement and Policy Implications | Winograd, G. | Discrete-time survival analysis | Higher education equity | Students with overlapping disadvantaged backgrounds face lower graduation probabilities | Demonstrates utility of intersectional analysis in equity research |
| Prospective assessment of learning curve and impact of intensive versus progressive training in colonoscopy competence among French residents | Wintzer-Wehekind, L.; Moulis, L.; et al. | Learning Curve-CUSUM analysis | Medical training | 21% of residents achieved colonoscopy competency after 204 procedures. | Not relevant (medical focus). |
| Association between SARS-CoV-2 viral load and serum biomarkers with mortality in Mexican patients | Razo-Blanco-Hernández, D. M.; Hernández-Mariano, J. A.; et al. | Cox regression | Public health | Higher viral load and serum biomarkers (e.g., BUN, glucose) predict COVID-19 mortality. | Not relevant (public health focus). |
| Staying Put: Positive Spillovers on Teacher Retention From a Middle School Science Initiative | Shiferaw, M.; O’Hagan, K. G.; Weinstein, M. | Discrete-time hazard model | K-12 workforce | UA program reduces teacher turnover by 3.8 percentage points. | Links professional development to retention; aligns with attrition literature. |
| Tech Equity: A Survival Analysis of an Undergraduate Computer Science Supplemental Education Program | Creps, R.; Islem, S.; et al. | Survival analysis | Higher education (CS) | Racial/gender disparities persist, but credit-bearing courses improve completion. | Highlights equity gaps in STEM interventions. |
| Is Admission Enough? University Persistence of Students Admitted Through Affirmative Action Policies in Chile | Corradi, B.; Espinosa, D.; et al. | Multilevel discrete-time survival model | Higher education equity | Affirmative action admits face higher dropout, independent of grades. | Critical for equity policy evaluation. |
| A longitudinal study of Head Start teacher turnover trends and factors | Kwon, K. A.; Jang, W.; et al. | Multilevel survival analysis | Early childhood education | Higher education and depressive symptoms increase turnover risk. | Extends survival analysis to early childhood workforce. |
| Fleeing School Choice? Resident Student Exit from Suburban School Districts | Lenhoff, S. W.; Pogodzinski, B.; et al. | Multilevel discrete-time survival analysis | K-12 policy | Rising Black/nonresident enrollment predicts resident student exit via choice. | Reveals racialized dynamics of school choice policies. |

### Appendix B: Regular Expressions Used for Method and Success Keyword Detection

|  |  |
| --- | --- |
| **Method** | **Search Terms / Patterns** |
| Kaplan-Meier | "kaplan[- ]meier", "km curve" |
| Cox Proportional Hazards | "cox proportional hazards", "cox regression", "cox model" |
| Discrete-Time Survival | "discrete[- ]time survival", "logistic hazard" |
| Competing Risks | "competing risks", "cause[- ]specific hazard" |
| Event History Analysis | "event history analysis" |
| Hazard Model | "hazard model", "hazard function" |
| Multilevel Survival | "multilevel survival", "hierarchical survival" |
| Machine Learning Survival | "random survival forest", "survival tree", "deep survival", "machine learning.\*survival" |
| Time-varying Covariates | "time[- ]varying covariates", "time[- ]dependent covariates" |
| Parametric Models | "weibull model", "exponential survival", "parametric survival" |
| Dyadic Survival | "dyadic survival" |

Success keywords: significant, improve, accurate, robust, validated, predictive power, fit well, low error.