

# Massive Open Online Courses (MOOCs) Dropout Rate in the World: A Protocol for Systematic Review and Meta-analysis

Manoosh Mehrabi<sup>1\*</sup>, PhD;  Ali Reza Safarpour<sup>2</sup>, PhD; Abbas Ali Keshtkar<sup>3</sup>, PhD

<sup>1</sup>Department of e-Learning in Medical Sciences, Virtual School, Center of Excellence for e-Learning in Medical Sciences, Shiraz University of Medical Sciences, Shiraz, Iran

<sup>2</sup>Gastroenterohepatology Research Center, Shiraz University of Medical Sciences, Shiraz, Iran.

<sup>3</sup>Department of Health Sciences Education Development, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

## ABSTRACT

**Introduction:** Massive open online course (MOOC) is an online course that is open, meaning there are no barriers to entry, and entails no special educational costs or features. Recently, MOOCs have received increasing popularity throughout the world. Regardless of the subject taught and the university providing the course, the dropout rate of MOOCs is one of the most important challenges ahead. The objective of this systematic review is to estimate the global rate of MOOCs dropout and factors affecting this frequency.

**Methods:** This systematic review will search MEDLINE/PubMed, Scopus, Web of Science (Clarivate Analytics), Embase (Embase.com), ASSIA, CINAHL, Education Research, BEI, and Eric databases systematically according to predefined criteria without language restrictions to retrieve prospective and retrospective observational studies conducted between the 1<sup>st</sup> of January 2000 and 31<sup>th</sup> of December 2021, evaluating the frequency of leaving MOOCs throughout the world. Discordances between the two different authors through the processes of screening, selection, quality assessment, and data extraction will be settled via discussion and if the issue cannot be resolved, a third expert advice will be requested. For all studies, forest plots will be shown to represent the separate and pooled frequency along with their 95% confidence intervals. To examine statistical heterogeneity, the Q-statistic test and the I<sup>2</sup> statistic will be utilized. To investigate potential reporting bias and non-significant study effects, funnel plots will be employed. Tests, such as Begg's and Egger's will also be carried out. The time trends for MOOCs dropout rate will be calculated using a cumulative meta-analysis.

**Conclusion:** As dropout rate is one of the most challenges that universities may encounter, this systematic review will help universities extend their view, save their resources, or maybe design their MOOCs differently. This protocol is registered in Open Science Framework (OSF), available at: <https://osf.io/jgyqx/>

**Keywords:** MOOCs, Dropout rate, Systematic review protocol

*\*Corresponding author:*  
Manoosh Mehrabi, PhD;  
Department of e-Learning  
in Medical Sciences, Virtual  
School, Center of Excellence  
for e-Learning in Medical  
Sciences, Shiraz University of  
Medical Sciences, Shiraz, Iran  
Tel: +98 9173069822  
Email: mehrabi.manoosh@  
gmail.com

*Please cite this paper as:*  
Mehrabi M, Safarpour AR,  
Keshtkar AA. Massive Open  
Online Courses (MOOCs)  
Dropout Rate in the World: A  
Protocol for Systematic Review  
and Meta-analysis. *Interdiscip  
J Virtual Learn Med Sci.*  
2022;13(2):85-92. doi:10.30476/  
IJVLM.2022.94572.1138.

Received: 30-02-2022

Revised: 17-03-2022

Accepted: 20-04-2022

## Introduction

Massive open online courses (MOOCs) have enjoyed increasing popularity in the world for the last 10 years (1-6) and with the growth of the Internet and educational technologies, the number of people joining MOOCs is increasing every day (1, 6, 7). According to the European open training report, providing MOOC-related statistics in European countries, there was a 130% increase in the number of MOOCs from September 2014 to September 2015 (8).

It appears that the need for MOOCs is apparent to all universities and governments and currently, these projects are encouraged and receive funding (1).

By definition, an MOOC is a course that is open, meaning there are no barriers to entry, and entails no special educational costs or features. It is provided online through the Internet and many people can connect to its educational environment (4).

In practice, the dropout rate of MOOCs, regardless of the subject taught and the university providing the course, is among the challenges MOOCs face (1-10). Online course dropout hinders universities from meeting the minimum expectations of such education (7). Course dropout has also been a concern to educational institutions because it is a waste of social, academic, and economic resources when a student begins a course but does not finish it (11). This dropout is caused by too many factors, such as academic skills of the students, their prior experiences, course design, feedback, and others (12).

Evaluating the global status of course dropout can provide crucial information that will help save resources and design new generations of MOOCs (9). The present systematic review addresses the gap in the literature on the MOOC dropout rate in the world.

According to our knowledge, in a study in 2017, Josksmovic et al. conducted a systematic review of approaches to model learning in MOOCs and reviewed studies on the prediction and measurement of students' staying in MOOCs. As one of the objectives of the analysis, the number of enrollments and

completions of courses was also addressed. Among the 13 studies that mentioned the exact number of enrollments and completions of courses, the highest and lowest dropout rates were 98% and 53.3%, respectively (9).

Furthermore, three review studies have reported related findings as follows:

A study by Paton showed that the dropout rate was 65% among Vocational Education and Training (VET) learners, 71% in Australian universities, and 72% in international universities. This study mentioned poor transferability and the narrow timeframe as its limitations (13).

Another study by Bezerra did not mainly aim at estimating course dropout rates, but mentioned a preliminary study that examined 91 courses in 2014 and found a dropout rate of 90% (8).

The third review study examined the causes of MOOC dropout and solutions for reducing it. The authors examined a study by Jordan (2013), which reported a course dropout rate of more than 90%. They also referred to a study by Meyer (2012), reporting an 80%-90% dropout rate with only 2% of the enrollments culminating in receiving a certificate. The authors also referred to other research in the field and reported a course dropout rate of 88.5% for a machine-learning course by Coursera, 94% for a bioelectricity course by Duke University, and 98.5% for a pattern-oriented software architecture course (14).

In comparison to previous studies, this systematic review will be a more precise and comprehensive search (all related electronic databases, grey literature, and relevant internet resources) with a longer time interval to estimate the rate of dropout of online courses and its affected factors throughout the world.

The current work aimed to combine the data from studies conducted since 2000 (MOOCs were introduced in 2008), identified through searching various databases and gray literature with no language restrictions, to estimate the global rate of MOOCs dropout systematically after comprehensive evaluation of the studies.

## Objectives

### Primary Objective

The primary objective of this systematic review is to estimate the accurate global rate of MOOCs dropout throughout the world between 2000 and 2021.

### Secondary Objectives

1. Estimation of the global frequency of MOOC dropout with age group, gender, and job type.
2. Estimation of the global frequency of MOOC dropout through its geographical distribution around the world.
3. Estimation of the global frequency of MOOC dropout via course duration.
4. Estimation of the global frequency of MOOC dropout using the level of the courses, including courses at the bachelor, master, and Ph.D. levels.
5. Estimation of the global frequency of MOOC dropout with course type, including medical courses versus non-medical courses.
6. Estimation of the global frequency of MOOC dropout using the ranks of universities.
7. Determination of the temporal trend of the global frequency of MOOC dropout over the last two decades and the trend before and after Covid-19 pandemic.
8. Assessment of the potential heterogeneity in the global frequency of MOOCs dropout in the world and finding its possible causes.

## Methods

This study will be a prevalence type of systematic reviews. The Centre for Reviews and Dissemination (CRD) standards were used to create this procedure (15) and it will be published using the Meta-analysis of Observational Studies in Epidemiology (MOOSE) recommendations (16). The studies will be chosen in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis-Protocols (PRISMA-P) 2015 guidelines (17). Based on the study question in the category of prevalence (frequency), the population is the students participating in MOOCs courses,

the outcome is the dropout rate, location is all around the world, our indicator is frequency, and the study design is a systematic review.

### Eligibility Criteria of Primary Studies

#### Inclusion Criteria

**Type of the studies:** This systematic review will include prospective and retrospective observational studies (cohort, case-control and cross-sectional studies) that evaluate the frequency of leaving MOOCs throughout the world. The studies should contain the frequency (%) of the people who left courses. No restriction on sample size and language will be applied for primary studies.

**Type of participants:** All the students who have participated in an MOOC, with ages equal or above 18 years (adult), in either gender (male and female), any races or ethnicity, in all geographical regions, and with all job types will be included in the current research.

#### Exclusion Criteria

Studies will be excluded from the systematic review if included in the following categories: the studies carried out before 01.01.2000 and after 31.12.2021; letters, editorials, case reports and commentaries; duplicates which are published in multiple journals (we will use the most recent and most inclusive version); and studies without any data of the rate of MOOC dropout.

## Outcomes

### Primary Outcome

Evaluation of the rate of MOOCs dropout during online courses in related studies is our study's primary outcome.

### Secondary Outcomes

Assessment of the factors possibly affecting the rate of dropout in MOOC courses is the secondary outcome in this review. These factors include gender, age, the jobs of the participants, setting of online courses, length of online courses, level (bachelor, master, PhD) of online course, type of the courses (medical vs non-medical), and University

rank of the course organizer.

### *Study Setting*

The systematic review would not be constrained by primary study's geographical location. All the available studies conducted to estimate the dropout rate of MOOCs around the world will be included.

### *Search Strategy and Literature Sources*

#### *Search Strategy Components*

We will search MEDLINE/PubMed, Scopus, Web of Science (Clarivate Analytics), Embase (Embase.com), ASSIA, CINAHL, Education Research, BEI, ERIC (Educational Resource Information Center) from 01.01.2000 to 31.12.2021. This systematic review will be conducted without any language restriction.

To reach the maximum inclusive search, the search strategy will only be based on the primary outcome component ("Massive open online courses" AND Dropout). To find the synonyms of the component, thesaurus systems, containing Emtree and MeSH, the free text method, the opinions of specialists and related articles and abstracts will be employed. The other approaches to be used for finding relevant studies include the following.

#### *Search in Key Journals and the Reference Lists of Last Included Studies*

Two main journals will be subjected to a manual issue-by-issue scan. The journals will be chosen based on an examination of the database search results and a search will be performed for journals that provide the largest pool of sources obtainable on the research topic, as determined by the study inclusion and exclusion criteria. A manual search of the reference library will also be carried out. A manual search of the reference lists of the articles chosen as final candidates for quality evaluation will also be performed.

### *Grey Literature*

In addition to contacting experts to obtain information on their relevant unpublished studies and related conferences, electronic databases, such as ProQuest and Scopus,

will be utilized to locate the relevance to the study area.

Electronic databases will be searched manually to collect appropriate conference papers and proceedings.

### *Search of Relevant Internet Resources*

To access more data, we will conduct a search of Class Central (<http://classcentral.com>).

### *Time Interval of Search*

All studies conducted between the 1<sup>st</sup> of January 2000 and 31<sup>th</sup> of December 2021 will be included.

To perform the most inclusive search, the search components include the dropout rate and MOOCs, as shown in Table 1.

This search approach will also work with other electronic databases.

All the stages of the search will be meticulously documented and presented with the final report. Endnote will keep track of all the searches in different databases.

### *Study Selection*

After the search strategy has been completed and tested by a colleague outside the authors' team, one of the authors will extract the relevant studies from the specified databases based on the inclusion and exclusion criteria. All the studies obtained will be stored in Endnote software for later analysis. After the searching process, two authors (ARS and MM) will review the title and abstract of the studies according to the inclusion and exclusion criteria and will find and extract the relevant studies. Final decision will be made according to their full text in the subsequent stage. Any discordance in all the steps will be decided with consensus and if the discrepancy is not resolved, the view of a third expert will be applied to determination the case.

### *Risk of Bias Assessment*

Two independent authors will use the ten-item tool, Hoy's assessment risk of bias in prevalence studies, to evaluate the



**Table 1:** Search syntax for Scopus

Number	Search Terms
1	((ALL(Dropout) AND ALL(Student)) OR (ALL(Dropouts) AND ALL(Student)) OR ALL("Student Dropout") OR ALL("School Dropouts") OR (ALL(Dropout) AND ALL(School)) OR (ALL(Dropouts) AND ALL(School)) OR ALL("School Dropout") OR (ALL(Dropouts) AND ALL(Education)) OR ALL("education dropout") OR ALL("education dropout") OR ALL("school dropout") OR ALL("school dropouts") OR ALL("student dropout") OR ALL("student dropouts") OR ALL("student dropouts") OR ALL("completion rate") OR ALL("completion rates") OR ALL("dropout rate") OR ALL("dropout rate") OR ALL("dropout rates") OR ALL("dropout rates") OR ALL("Participant engagement") OR ALL("learner engagement") OR ALL("course engagement") OR ALL("student engagement") OR ALL("attrition rate") OR ALL("attrition rates") OR ALL("learner participation") OR ALL("course participation") OR ALL("student participation") OR ALL("non completion rate") OR ALL("non completion rates") OR ALL("educational status") OR ALL(Dropout*) OR ALL("student withdraw") OR ALL("learner withdraw") OR ALL("course withdraw") OR ALL("participant withdraw") OR ALL("student abandon") OR ALL("learner abandon") OR ALL("course abandon") OR ALL("participant abandon") OR ALL("student discontinue") OR ALL("learner discontinue") OR ALL("course discontinue") OR ALL("participant discontinue") OR ALL("student leave") OR ALL("learner leave") OR ALL("course leave") OR ALL("participant leave") OR ALL("student quit") OR ALL("learner quit") OR ALL("course quit") OR ALL("participant quit") OR ALL("student achievement") OR ALL("learner achievement") OR ALL("course achievement") OR ALL("participant achievement") OR ALL("student competence") OR ALL("learner competence") OR ALL("participant competence") OR ALL("course competence") OR ALL("student performance") OR ALL("learner performance") OR ALL("course performance") OR ALL("participant performance") OR ALL("student progression") OR ALL("learner progression") OR ALL("course progression") OR ALL("participant progression") OR ALL("retention rate") OR ALL(Dropout*) OR ALL(Withdraw*) OR ALL(abandon) OR ALL(discontinue) OR ALL(leave) OR ALL(quit) OR ALL(achievement*) OR ALL(competence*) OR ALL(perform*) OR ALL(progress*))
2	(ALL(MOOCs) OR ALL("Massive Open Online Courses") OR ALL(MOOC))
3	1 AND 2
4	PUBYEAR AFT 2000
5	3 AND 4

methodological quality of primary studies (18). This tool contains the subsequent items: items one to four evaluate external validity and comprise the representativeness of the target population, the representativeness of the sample population (sampling frame), random selection, and the non-response bias. Items five to ten measure internal validity and contain data gathering from the subjects or substitutions, acceptable case definition, reliability and validity of the measurement instruments, same mode of data collection used for all the subjects, suitability of the length of the shortest prevalence period, and the appropriateness of the numerator (s) and denominator (s). Two authors will

apply the standards and argue about discordance, which will be fixed in

consultation with a third expert author.

#### *Data Extraction*

For the final included studies, two authors will independently extract the following summary data according to a standard developed and tested form. Any disagreement between the two authors in this step will be resolved through discussion and opinion of expert person in non-resolved cases. We will extract the following key data: Author/s and year of the study; Country of conducting study; Affiliation of first author; Type of participant/study population/demographic, Characteristics, including age, gender, and job; Sample size; Type of study (study design); Type of outcomes measured (primary or secondary); Length of online course

(duration); Level of the course, including bachelor, master, and PhD; Type of the course, namely medical and non-medical courses;

Rank of the University; and Findings/results, including dropout rate as frequency (%) in online course and causes of dropout.

### *Data Analysis and Synthesis*

This review will determine the frequency of the dropout rate in online courses. This indicator will be calculated from the following formula; *The number of students who dropped out of the course/sample size in that study*. The results of each included study will be presented as a frequency (%) of MOOCs dropout and 95% confidence interval (CI). We will show the extracted data for each study in Table 1 in the final report.

All the data (the main assessed indicator and frequency) will be pooled from final included research in a meta-analysis using the random-effects model if there is no significant methodological heterogeneity, inconsistency, or a high risk of bias. If meta-analysis is not possible due to severe methodological heterogeneity and variation between studies, we will not combine the results and a narrative qualitative discussion based on findings from individual studies will take place. Forest plots will be plotted for all the studies to demonstrate the separated and pooled frequencies and their corresponding 95% CIs. The software used in the present study will be Stata V.14 (Stata Corp).

### *Assessment of Heterogeneity*

The Q-statistic test and  $I^2$  statistic and their corresponding 95% CIs will be employed to assess the statistical heterogeneity of the frequency values in the included works. The references according to Cochrane Handbook (Heterogeneity values of 0%–40% will take place as perhaps not important; Heterogeneity values of 30%–60% as moderate heterogeneity; Heterogeneity values of 50%–90% as substantial heterogeneity; and Heterogeneity values of 75%–100% as considerable heterogeneity) will be used as the bases to determine the degree of heterogeneity.

The level of statistical significance will be set at  $P < 0.05$  for the Q-test.

### *Sub Group Analysis*

The purpose of the subgroup analysis in this study is to determine if varying variables have an effect on the MOOCs dropout rate and in what context this occurs. Sub-group analysis or meta-regression, if sufficient data are available, will be used appropriately to investigate the effect of statistical heterogeneity. Herein, variables, such as age, sex, duration of the course, and types (medical and non-medical), and degree level of the courses, will be utilized in sub group analysis.

### *Sensitivity Analysis*

We will perform various sensitivity analyses to assess methodological quality (risk of bias assessment), design limitations, data analysis considerations, sample size, and effect of missing data. Sensitivity analysis also will be used to determine how excluded studies could have influenced the overall result. The one-out remove method will also be applied for sensitivity analysis. If one of the combinations (K-1) of the studies exhibits a different result from the others, we will carefully consider the features of that study.

### *Assessment of Publication Bias*

The first approach for dealing with publishing bias is to conduct the most comprehensive search possible at the study's search stage. If there are sufficient studies, both Funnel Plot and Begg's and Egger's statistical tests will be run to evaluate publication bias. The 'trim and fill' approach would be used if significant findings ( $p > 0.1$ ) show a publishing bias.

## **Discussion**

The pooled rate of MOOC dropout around the world will be estimated in this systematic review and meta-analysis study. It will also provide evidence of the causes for the probable variation in the reported prevalence of this type of online education.

Dropout rate of MOOCs is one of the most significant challenges that universities may encounter. On the other hand, the number of people joining MOOCs is increasing. Moreover, it is apparent that these projects need investment both in human resources and economic resources. Since this study will use comprehensive and careful methods in all the phases of the systematic review and meta-analysis, the evidences achieved will be wholly reliable.

### *Implications to Practice*

This systematic review will help universities deciding on launching these kinds of platforms to extend their view and correct their expectations about staying in the course. This study will also provide information which will help universities save their resources or maybe design their MOOCs differently.

### **Authors' Contributions**

MM: Idea and conception, design the work, drafted the work, approved submitted version. ARS: Design the work, drafted the work, approved submitted version, guarantor. AAK: Revision of the work, approved submitted version.

**Conflict of Interest:** None declared.

### **Funding**

The authors received no funding.

### **References**

- 1 Gregori EB, Zhang J, Galván-Fernández C, de Asís Fernández-Navarro F. Learner support in MOOCs: Identifying variables linked to completion. *Computers & Education*. 2018;122:153-68.
- 2 Tsai Y-h, Lin C-h, Hong J-c, Tai K-h. The effects of metacognition on online learning interest and continuance to learn with MOOCs. *Computers & Education*. 2018;121:18-29.
- 3 Jordan K. Initial trends in enrolment and completion of massive open online courses. *International Review of Research in Open and Distributed Learning*. 2014;15(1):133-60.
- 4 Xu B, Yang D. Study partners recommendation for xMOOCs learners. *Computational intelligence and neuroscience*. 2015;2015.
- 5 Eriksson T, Adawi T, Stöhr C. "Time is the bottleneck": a qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education*. 2017;29(1):133-46.
- 6 Gomez-Zermeno MG, de La Garza LA. Research analysis on MOOC course dropout and retention rates. *Turkish online journal of distance education*. 2016;17(2).
- 7 Gregori P, Martínez V, Moyano-Fernández JJ. Basic actions to reduce dropout rates in distance learning. *Evaluation and program planning*. 2018;66:48-52.
- 8 BEZERRA LN, SILVA MT. A review of literature on the reasons that cause the high dropout rates in the MOOCs. *Revista Espacios*. 2017;38(05).
- 9 Joksimović S, Poquet O, Kovanović V, Dowell N, Mills C, Gašević D, et al. How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*. 2018;88(1):43-86.
- 10 Goldberg LR, Bell E, King C, O'Mara C, McInerney F, Robinson A, et al. Relationship between participants' level of education and engagement in their completion of the Understanding Dementia Massive Open Online Course. *BMC medical education*. 2015;15(1):1-7.
- 11 de Castro e Lima Baesse D, Grisolia AM, de Oliveira AEF. Pedagogical monitoring as a tool to reduce dropout in distance learning in family health. *BMC Medical Education*. 2016;16(1):1-8.
- 12 Aldowah H, Al-Samarraie H, Alzahrani AI, Alalwan N. Factors affecting student dropout in MOOCs: a cause and effect decision-making model. *Journal of Computing in Higher Education*. 2020;32(2):429-54.
- 13 Paton RM, Fluck AE, Scanlan JD. Engagement and retention in VET MOOCs

- and online courses: A systematic review of literature from 2013 to 2017. *Computers & Education*. 2018;125:191-201.
- 14 Khalil H, Ebner M. MOOCs completion rates and possible methods to improve retention-A literature review. *EdMedia+ innovate learning*. 2014:1305-13.
  - 15 Dissemination C. *Systematic reviews: CRD's guidance for undertaking reviews in healthcare*. York: University of York NHS Centre for Reviews & Dissemination. 2009.
  - 16 Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *Jama*. 2000;283(15):2008-12.
  - 17 Moher D, Liberati A, Tetzlaff J, Altman DG, Group\* P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*. 2009;151(4):264-9.
  - 18 Hoy D, Brooks P, Woolf A, Blyth F, March L, Bain C, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. *Journal of clinical epidemiology*. 2012;65(9):934-9.