The value of long noncoding RNAs for predicting the recurrence of endometriosis
A protocol for meta-analysis and bioinformatics analysis
Yihong Chen, MDa,b, Xinghui Liu, MDa,b, Lei He, MDa,b,∗

Abstract
Background: As a gynecological disease, endometriosis (EM) seriously endangers the health of women at the age of childbearing and is closely related to long noncoding RNAs (lncRNAs). Current studies have discovered that there are differential expressions of many kinds of lncRNAs in EM. However, whether lncRNAs can be applied as a new marker for the prediction of the recurrence of EM is still controversial. In this study, meta-analysis and bioinformatics analysis were carried out to explore the value of lncRNAs as a predictor of the recurrence of EM and to analyze its biological role.

Methods: PubMed, Embase, and Web of Science databases were searched through computer and the articles published from the self-built database to April 2021 were collected. According to the inclusion and exclusion criteria, the literature was screened, and the quality of the inclusion study was evaluated. Stata 16.0 software was used for meta-analysis. The co-expression genes related to IncRNAs were screened by online tool Co-LncRNA. Then David for Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analysis were conducted. A competitive endogenous RNA network that may exist in lncRNAs through Starbase was built.

Results: The results of this meta-analysis would be submitted to peer-reviewed journals for publication.

Conclusion: This meta-analysis could provide high-quality evidence support for lncRNAs, so as to predict the recurrence of EM. At the same time, we use bioinformatics technology to predict and analyze its biological effects, which provides a theoretical basis for further experimental verification.

Ethics and dissemination: The private information from individuals will not be published. This systematic review also should not damage participants’ rights. Ethical approval is not available. The results may be published in a peer-reviewed journal or disseminated in relevant conferences.

OSF Registration Number: DOI 10.17605/OSF.IO/MF3QJ.

Abbreviations: EM = endometriosis, lncRNAs = long noncoding RNAs.

Keywords: bioinformatics, endometriosis, long noncoding RNAs, meta-analysis, protocol, recurrence

1. Introduction
Characterized by the appearance of endometrial matrix and glands outside the uterine cavity,[1] endometriosis (EM) is a gynecological disease. The pathogenesis of EM is multifactorial, and hormone, immune, environmental and genetic factors are all related to it.[2,3] Although the exact pathogenesis is still unclear, genetic factors, especially epigenetic factors, play an important role.[4] After treatment, EM has a high recurrence rate, and no effective recurrence predictors could be found. To this end, finding noninvasive biomarkers for EM has been an ongoing and challenging problem.

As an important branch of epigenetic research, noncoding RNA can regulate the expression of related genes, then play a significant role in abnormal proliferation, apoptosis, invasion, steroid hormone metabolism and its receptor expression, epithelial-mesenchymal transformation, angiogenesis and other processes of endometrial cells, and promote the occurrence and development of EM.[5] An increasing number of evidences indicate that long noncoding RNAs (lncRNAs) are involved in the genesis and development of EM in various ways.[5–11] In
addition, the potential of lncRNAs in improving the early diagnosis, evaluation, and treatment of EM has been widely explored.\textsuperscript{12}

In recent years, related studies have proved that lncRNAs have important clinical significance in predicting the recurrence of EM.\textsuperscript{13–16} However, the results of each study are inconsistent. For this reason, this study performed a meta-analysis to evaluate the accuracy of lncRNAs in predicting the recurrence of EM, and to provide evidence support for the application of clinical noninvasive markers. In this study, in order to further understand the biological role of lncRNAs in EM, a competitive endogenous RNA network was constructed by means of bioinformatics to reveal the expression and function of regulatory genes, which provide a theoretical basis for further experimental verification.

2. Methods

2.1. Study registration

This meta-analysis protocol is based on the Preferred Reporting Items for Systematic Reviews and Meta-analysis Protocols (PRISMA-P) statement guidelines.\textsuperscript{17} The protocol of the systematic review was registered on Open Science Framework, and the registration number is DOI 10.17605/OSF.IO/MF3QJ.

2.2. Data sources and retrieval strategy

We searched the Web of Science, PubMed, and EMBASE databases to identify all potentially eligible articles from inception to April 2021. The detailed search strategies are listed in Table 1.

2.3. Inclusion criteria for study selection

2.3.1. Inclusion criteria.

1. The patients who were diagnosed with EM. At the same time, surgical treatment was performed.
2. The samples of lncRNA come from tissue, platelets, serum, peripheral blood mononuclear cells, plasma, and whole blood.
3. To explore the diagnostic value of lncRNAs on the diagnosis of recurrence of EM.
4. True positive, false positive, false negative, and true negative can be calculated based on the information in the literature.

2.3.2. Exclusion criteria.

1. Repeatedly published research.
2. Animal experiment.
3. Comments, case reports, conference summaries, and meta-analysis.
4. Insufficient data.

2.4. Data collection and analysis

The literature screening process is displayed in Figure 1. According to the unified inclusion and exclusion criteria, after the preliminary screening of the literature, the data of the included literature were extracted independently by 2 researchers and cross-checked to confirm whether the data were accurate or not. The data collected include the first author, the number of years published, the nationality of the study, the design of the study, the source of the sample, the detection method, the type of lncRNA, the longest follow-up period, outcome indicators, etc.

Furthermore, in view of the fact that some studies only provide receiver operating characteristic curve, it is necessary to use Engauge Digitizer4.1 version to extract true positive, false positive, false negative, and true negative.\textsuperscript{18,19}

2.5. Quality assessment

The risk of bias in the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) score system.\textsuperscript{20}

2.6. Outcome indicator

The outcomes include pooled sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, area under the curve, and their 95% confidence intervals.

2.7. Management of missing data

If there exists insufficient or missing data in the literature, we would only analyze the currently available data and discuss its potential value.
2.8. Statistical analysis

All statistical analyses were conducted using Stata 16.0 (STATA Corp, College Station, TX) software. We calculated the pooled sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, and their 95% confidence intervals. What is more, the pooled diagnostic value of lncRNAs through the summary receiver operated characteristic curve and area under the curve was tested. The threshold effects were detected by using Spearman correlation coefficient. The calculation of heterogeneity was caused by the nonthreshold effects of Cochrane-Q and $I^2$ values, and a fixed-effect model (without obvious inhomogeneity) or a random-effects model (with significant heterogeneity) was employed to merge the data. The statistical test level was $\alpha = 0.05$.

2.9. Additional analysis

2.9.1. Subgroup analysis. According to the detection methods of lncRNAs, ethnicity, and the source of lncRNAs, we analyzed the subgroup.

2.9.2. Sensitivity analysis. Sensitivity analysis was performed via sequential deletion of a single included study to test.

2.9.3. Reporting bias. The publication bias was determined by carrying out Deeks funnel plot asymmetry test.
3. Bioinformatics analysis

3.1. Screening of genes related to the role of IncRNAs

The genes that may be related to the expression level of IncRNAs were screened by online tool Co-LncRNA (http://www.biobigdata.com/Co-LncRNA/), and the intensity of co-expression was expressed by interaction score.[21]

3.2. Enrichment analysis of Gene Ontology biological function and Kyoto Encyclopedia of Genes and Genomes signaling pathway

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analyses were performed on the screened co-expressed genes using the online tool David (https://david.ncifcrf.gov/).[22]

3.3. Construction of competitive endogenous RNA network of IncRNAs

The online tool starbase (http://starbase.sysu.edu.cn/index.php) was used to study the micro RNA molecules targeted by IncRNAs and downstream target mRNA molecules.

4. Ethics and dissemination

The content of this article does not involve moral approval or ethical review and would be presented in print or at relevant conferences.

5. Discussion

As an important member of the noncoding RNA family in vivo, IncRNAs play an important regulatory role in the occurrence and development of many diseases.[23,24] In recent years, with the development of high-throughput sequencing technology, more and more attention has been paid to the important functions of IncRNAs. As a common disease in women at the age of childbearing, EM seriously affects their physiological health.[25,26] Although the cause of the disease is not clear, genetic factors play an important role in it. Although scholars have revealed that a variety of differentially expressed IncRNAs may play an important role in the occurrence and development of EMs.[27] Existing studies have indicated that IncRNAs promote the occurrence and development of EM through proliferation, invasion, metastasis, apoptosis, autophagy, and so on.[9,28–30] Further screening of IncRNAs, with high specificity and sensitivity, has important clinical significance for the diagnosis, treatment, and prognosis of EM. In this study, meta-analysis and bioinformatics were used to screen IncRNA, with high specificity and sensitivity for predicting EM recurrence, thus constructing a competitive endogenous RNA network to reveal the occurrence and development of EM.

Author contributions

Conceptualization: Lei He, Yihong Chen.
Data curation: Lei He, Yihong Chen.
Formal analysis: Xinghui Liu and Yihong Chen.
Funding acquisition: Lei He.
Methodology: Yihong Chen.
Project administration: Lei He.
Software: Xinghui Liu, Yihong Chen.

Supervision: Lei He.
Validation: Yihong Chen.
Visualization: Xinghui Liu, Yihong Chen.
Writing – original draft: Lei He, Yihong Chen.
Writing – review & editing: Lei He, Yihong Chen.

References

[1] Giudice LC, Kao LC. Endometriosis. Lancet 2004;364:1789–99.
[2] Parazzini F, Espostio G, Tozzi L, Noli S, Bianchi S, et al. Epidemiology of endometriosis and its comorbidities. Eur J Obster Gynecol Reprod Biol 2017;209:3–7.
[3] Shafrir AL, Farland LV, Shah DK, et al. Risk for and consequences of endometriosis: a critical epidemiologic review. Best Practice Res Clin Obster Gynaecol 2018;51:1–15.
[4] Králiková M, Lozan P, Verticka V. Endometriosis and cancer. Women Health 2014;10:591–7.
[5] Josep MA, Dolors SI, Juan GE, Barceló-Molina M, Braza-Boils A, Sandoval J, et al. miRNAs regulation and its role as biomarkers in endometriosis. Int J Mol Sci 2016;17:93.
[6] Wang H, Sha L, Huang L, et al. LINC00261 functions as a competing endogenous RNA to regulate BCL2L11 expression by sponging miR-132-3p in endometriosis. Am J Transl Res 2019;11:2269–79.
[7] Liu J, Wang Y, Chen P, et al. AC002454.1 and CDK6 synergistically promote endometrial cell migration and invasion in endometriosis. Reproduction 2019;157:535–43.
[8] Lin D, Huang Q, Wu R, et al. Long non-coding RNA AFAPI1-S1 promoting epithelial-mesenchymal transition of endometriosis is correlated with transcription factor ZEB1. Am J Reprod Immunol 2019;81:13074.
[9] Liu H, Zhang Z, Xiong W, et al. Long non-coding RNA MALAT1 mediates hypoxia-induced pro-survival autophagy of endometrial stromal cells in endometriosis. J Cell Mol Med 2019;23:439–52.
[10] Lin K, Zhan H, Ma J, et al. Silencing of SRA1 regulates ER expression and attenuates the growth of stromal cells in ovarian endometriosis. Reprod Sci 2017;24:836–43.
[11] Li Y, Liu YD, Chen SL, et al. Down-regulation of long non-coding RNA MALAT1 inhibits granulosa cell proliferation in endometriosis by up-regulating P21 via activation of the ERK/MAPK pathway. Mol Hum Reprod 2019;25:17–29.
[12] Painir K, Schjenken JE, Robertson SA, Hull ML, et al. Non-coding RNAs in endometriosis: a narrative review. Hum Reprod Update 2018;24:497–515.
[13] Liu S, Xin W, Tang X, Qiu J, Zhang Y, Hua K, et al. LncRNA H19 overexpression in endometriosis and its utility as a novel biomarker for predicting recurrence. Reprod Sci 2020;27:1687–97.
[14] Qiu J, Zhang X, Ding Y, et al. 1856 circulating exosomal long noncoding RNA-TC0101441 as a non-invasive biomarker for the prediction of endometriosis severity and recurrence. J Minim Invasive Gynecol 2019;26:S170–1.
[15] Qiu J, Hua K. Exosomal long non-coding RNA NONHAT076754 facilitates endometriosis invasion and predicts endometriosis recurrence. J Minim Invasive Gynecol 2019;26:S12.
[16] Huang H, Zhu Z, Song Y. Downregulation of lncrna uca1 as a diagnostic and prognostic biomarker for ovarian endometriosis. Rev Assoc Med Bras 2019;65:336–41.
[17] Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;350:g7647.
[18] Williamson PR, Smith CT, Hutton JL, Marson AG, et al. Aggregate data meta-analysis with time-to-event outcomes. Stat Med 2002;21:3373–51.
[19] Tierney JF, Stewart LA, Gherzi D, et al. Practical methods for incorporating summary time-to-event data into meta-analysis. Trials 2007;8:1–16.
[20] Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern M 2019;153:529–36.
[21] Zhao Z, Bai J, Wu A, et al. Co-LncRNA: investigating the IncRNA combinatorial effects in GO annotations and KEGG pathways based on human RNA-Seq data. Database 2015:2015:bav082.
[22] Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.
[23] Wen J, Xu J, Sun Q, Xing C, Yin W, et al. Upregulation of long non-coding RNA PCAT-1 contributes to cell proliferation, migration and apoptosis in hepatocellular carcinoma. Mol Med Rep 2016;13:4481–6.

[24] Xu W, Chang J, Du X, Hou J, et al. Long non-coding RNA PCAT-1 contributes to tumorigenesis by regulating FSCN1 via miR-145-5p in prostate cancer. Biomed Pharmacother 2017;95:1112–8.

[25] Rossi AC, Prefumo F. The effects of surgery for endometriosis on pregnancy outcomes following in vitro fertilization and embryo transfer: a systematic review and meta-analysis. Arch Gynecol Obstet 2016;294:647–55.

[26] Parkin KL, Fazleabas AT. Uterine leukocyte function and dysfunction: a hypothesis on the impact of endometriosis. Am J Reprod Immunol 2016;75:411–7.

[27] Zhu MB, Chen LP, Hu M, Shi Z, Liu YN, et al. Effects of lncRNA BANCR on endometriosis through ERK/MAPK pathway. Eur Rev Med Pharmacol Sci 2019;23:6806–12.

[28] Liu S, Qiu J, Tang X, Cui H, Zhang Q, Yang Q, et al. LncRNA-H19 regulates cell proliferation and invasion of ectopic endometrium by targeting ITGB3 via modulating miR-124-3p. Exp Cell Res 2019;381:215–22.

[29] Xu Z, Zhang L, Yu Q, Zhang Y, Yan L, Chen ZJ, et al. The estrogen-regulated lncRNA H19/miR-216a-5p axis alters stromal cell invasion and migration via ACTA2 in endometriosis. Mol Hum Reprod 2019;25:550–61.

[30] Sha L, Huang L, Luo X, et al. Long non-coding RNA LINC00261 inhibits cell growth and migration in endometriosis. J Obstet Gynaecol Res 2017;43:1563–9.