Available Online: 24 November 2026 DOI: 10.54254/2977-3903/2026.39241

AI-assisted siRNA drug design: potential applications in PCSK9-targeted lipid-lowering therapy

Siyuan Wang

School of Arts and Sciences, Rutgers University, New Brunswick, USA

Siyuanwang2006@163.com

Abstract. Some patients do not reach their low-density lipoprotein cholesterol targets despite existing lipid-lowering treatment. Small interfering RNA drugs targeting PCSK9 offer an additional treatment option. However, silencing activity varies among candidate sequences, and chemical modifications and delivery conditions also affect their therapeutic potential. This review evaluates the evidence supporting AI-assisted early screening of PCSK9-targeting siRNAs and the limits of its application. This paper examines studies of activity prediction for chemically modified siRNAs, computational evaluation using AttSiOff, clinical trials of inclisiran, and research on modification optimization and GalNAc-mediated liver-targeted delivery. Current evidence indicates that machine learning can use sequence information and selected modification features to help prioritize candidates, while inclisiran trials support the lipid-lowering effects of PCSK9-targeting siRNAs. These findings are not yet sufficient to establish that integrating AI tools improves actual screening efficiency or therapeutic efficacy. Future studies should first standardize modification and delivery conditions, then conduct prospective controlled screening using hit rate and experimental resources as evaluation measures, followed by validation of candidate efficacy and safety in vivo.

Keywords: AI, siRNA, PCSK9, hypercholesterolemia

1. Introduction

Hypercholesterolemia is an important modifiable risk factor for atherosclerotic cardiovascular disease. Although statins effectively lower Low-Density Lipoprotein Cholesterol (LDL-C), some patients still require additional lipid-lowering treatment despite receiving the maximum tolerated dose [1]. PCSK9 promotes the degradation of LDL receptors on hepatocytes, reducing the clearance of LDL from the blood [2]. Inclisiran, an siRNA drug targeting PCSK9, reduces PCSK9 synthesis through RNA interference, and thereby promotes LDL clearance. In ORION-10 and ORION-11, placebo-adjusted reductions in LDL-C at day 510 were 52.3% and 49.9%, respectively, supporting the clinical lipid-lowering effect of this therapeutic strategy [3].

However, siRNA candidates directed at the same target can differ in silencing activity, and chemical modifications and delivery conditions also affect their therapeutic potential[4-6]. Machine learning can use sequence and modification information to predict silencing activity. Tools such as AttSiOff combine activity prediction with screening for potential off-target binding sites to provide a computational basis for candidate

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

prioritization [4, 5]. Whether these predictions translate into improved efficiency in the actual screening of PCSK9-targeting siRNAs requires further evaluation.

This narrative review compares studies of siRNA activity prediction, computational evaluation using AttSiOff, and clinical evidence for inclisiran, with particular attention to their validation methods and scope of applicability. It also examines how chemical modifications and liver-targeted delivery affect candidate evaluation. By distinguishing computational predictions from experimental silencing results and clinical lipid-lowering evidence, the review considers the extent to which AI-assisted early screening of PCSK9-targeting siRNAs may improve efficiency and what further validation is needed.

2. Evidence for AI-assisted siRNA screening and PCSK9-targeted therapy

2.1. AI-based prediction of siRNA activity

Martinelli investigated whether machine learning could identify candidates with high silencing activity from siRNA nucleotide sequences and chemical modification patterns. The study selected 907 chemically modified siRNAs, each 21 nucleotides long, from the siRNAmod database. These records originated from 30 papers or patents and included sequences, modification information, and experimentally measured silencing activity [4].

The study used one-hot encoding to convert the nucleotide identities in the sense and antisense strands, together with the chemical modifications at each position, into numerical features. Three models were compared: a random forest, a multilayer perceptron neural network, and a linear support vector machine. Martinelli used silencing activity thresholds of 60%, 70%, and 80% to classify samples according to whether they met each threshold. The prediction task was therefore to determine whether a candidate would reach a specified activity threshold. The data were divided into training and test sets containing 75% and 25% of the samples, respectively, and model performance was assessed using five-fold cross-validation [4].

To further assess predictive performance, Martinelli compiled an external validation set of 57 siRNAs from six publications not included in the database, 53 of these siRNAs were chemically modified. At the 70% classification threshold, the areas under the receiver operating characteristic curve were 0.895 for the random forest, 0.850 for the neural network, and 0.887 for the support vector machine, which means the models could distinguish reasonably well between candidates that met the threshold and those that did not. Performance varied with the classification threshold and declined at the other thresholds, indicating that predictive performance remained dependent on how the task was defined [4].

Within this dataset of 907 modified siRNAs, sequence and modification information could be used together for activity classification. However, this finding remains limited by the modification types and activity thresholds examined in the study.

2.2. Clinical application of an siRNA drug

Inclisiran is an siRNA drug targeting PCSK9 mRNA. Its clinical lipid-lowering effects were evaluated in ORION-10 and ORION-11, two randomized, double-blind, placebo-controlled phase 3 trials. The trials enrolled 1,561 and 1,617 patients, respectively, whose LDL-C levels remained elevated despite treatment with the maximum tolerated statin dose. ORION-10 enrolled patients with established atherosclerotic cardiovascular disease, while ORION-11 also included patients with an equivalent level of cardiovascular risk. Patients were randomly assigned to receive subcutaneous inclisiran or placebo in addition to their existing lipid-lowering treatment. Injections were administered on days 1 and 90 and every six months thereafter, with a total observation period of 540 days. The primary efficacy endpoints included the placebo-adjusted

percentage change in LDL-C at day 510 and the placebo-adjusted time-adjusted average percentage change from day 90 to day 540 [3].

At day 510, the placebo-adjusted reductions in LDL-C were 52.3% in ORION-10 and 49.9% in ORION-11, and both comparisons were statistically significant. Overall adverse event rates were similar between the groups, although injection-site adverse events were more common with inclisiran [3]. Inclisiran provided an approximately 50% reduction in LDL-C in addition to maximally tolerated statin therapy, demonstrating that inhibition of PCSK9 synthesis can produce a clinical lipid-lowering effect. However, these results were obtained with a specific drug in defined patient populations and cannot be directly extrapolated to new candidates.

2.3. An existing AI-assisted siRNA design tool

Liu et al. developed AttSiOff. AttSiOff combines a self-attention-based model for predicting siRNA activity with modules for mRNA retrieval and off-target screening. After a target gene name is entered, the tool retrieves the corresponding mature mRNA sequence and generates candidate siRNAs using a sliding window. The pretrained RNA-FM model extracts sequence features from the siRNA and the mRNA surrounding its target region. The prediction model combines these features with thermodynamic parameters, secondary structure, and GC content to predict silencing activity and rank candidates. The off-target screening module identifies potential non-target binding sites to inform the assessment of candidate specificity [5].

To examine the tool's application to drug-related sequences, Liu et al. selected five existing siRNA drugs and retrospectively evaluated their underlying sequences after removing chemical modification information. The targets included PCSK9. Candidate sets were generated for each target, and the predicted rank of each drug sequence within its corresponding set was assessed, together with the number of potential off-target binding sites. All five sequences ranked within the top 11.2% of their respective candidate sets, indicating that the tool could assign relatively high ranks to these drug-related sequences [5]. Thus, AttSiOff could at least place the underlying sequence of a known effective PCSK9 siRNA among the higher-ranked candidates. However, because chemical modifications were excluded from the evaluation, a high rank does not establish that a sequence will retain high activity after modification.

2.4. Chemical modifications and delivery

Chemical modifications and delivery conditions must be considered when evaluating the therapeutic potential of siRNA candidates. Foster et al. optimized modification patterns by adjusting the positions and proportions of 2′-fluoro and 2′-O-methyl modifications in both siRNA strands, using in vitro screening and animal experiments. Some optimized designs retained in vitro silencing activity while improving the magnitude and duration of the in vivo effect. Further analyses suggested that these improvements were mainly attributable to increased metabolic stability. However, a fully 2′-O-methyl-modified control showed relatively high liver exposure but low Ago2 loading and no appreciable target gene silencing. These findings indicate that modification design must balance improved stability with preservation of RNA interference activity, efficacy cannot be inferred solely from tissue siRNA levels [7].

For delivery, Nair et al. conjugated siRNAs to the N-acetylgalactosamine ligand GalNAc to promote cellular uptake through the asialoglycoprotein receptor on hepatocytes. GalNAc conjugates showed greater uptake in primary mouse hepatocytes than unconjugated siRNAs. Both ligand competition and loss of a receptor subunit reduced uptake, supporting a receptor-mediated delivery mechanism. The study also demonstrated hepatic target gene silencing in mice, providing experimental evidence for GalNAc-mediated liver-targeted siRNA delivery [8]. Inclisiran likewise uses GalNAc conjugation to promote hepatocyte uptake

[1]. However, receptor-mediated uptake does not mean that siRNA has effectively reached the cytosol. Internalized siRNA must still be released from membrane-bound compartments such as endosomes before it can be loaded into the RNA-induced silencing complex and exert its effect in the cytosol. Brown et al. showed that endosomal escape is an important limiting step in effective GalNAc-siRNA delivery. Cellular uptake or total hepatic siRNA levels therefore cannot directly represent the amount available for gene silencing [9].

These studies also reveal a gap between AI-based activity predictions and candidate performance in vivo. Martinelli's model uses sequence information and modification types and positions as inputs, but does not directly predict metabolic stability, Ago2 loading, or liver-targeted delivery [4]. AttSiOff's retrospective evaluation of existing drugs excluded chemical modification information, so its rankings cannot represent the actual performance of the fully modified drugs [5]. Foster et al. further demonstrated that higher liver exposure does not necessarily produce stronger gene silencing [7]. When applying these tools to PCSK9 siRNA screening, their inputs and prediction scope must therefore be clearly defined. Candidate silencing activity and in vivo effects must then be experimentally validated under specified modification and delivery conditions.

3. Proposed applications and validation requirements for AI-assisted PCSK9 siRNA screening

Existing studies provide evidence at different levels for AI-assisted screening and PCSK9-targeted therapy. The three types of evidence address different questions: Martinelli's models assess discrimination in activity classification, AttSiOff evaluates the ranking of known drug sequences within candidate sets, and the ORION trials establish the lipid-lowering effects of inclisiran as a specific drug. Together, they support the idea of using AI tools to screen new PCSK9 siRNA candidates while drawing on inclisiran's target and delivery strategy, but they are not interchangeable. In a blinded, controlled experiment involving 14 genes, Wang et al. compared candidates selected by an SVM-assisted design workflow with those selected by the original algorithm. At a concentration of 3 nM, all 42 siRNAs selected by the SVM-assisted workflow met the criterion of at least an 80% reduction in target gene expression, giving an observed hit proportion of 42/42. More than 80% of the candidates selected by the original algorithm met the same criterion. This provides experimental support for the SVM-assisted workflow's ability to prioritize highly active siRNAs, but only for the candidates and conditions tested, it cannot be generalized to the screening hit rate for all siRNAs [10]. Moreover, the study did not target PCSK9 or evaluate the complete modification and delivery conditions required for clinical drugs. Whether its screening advantage extends to PCSK9 siRNA candidates therefore requires further testing.

These studies suggest a staged screening strategy that remains to be validated. AttSiOff could first predict activity and screen for potential off-target binding sites in candidate sequences targeting PCSK9. Chemical modification designs could then be assigned to the candidates, and post-modification silencing activity could be evaluated using an approach informed by Martinelli's study, provided that the training data cover the intended modification types [4, 5]. Importantly, AttSiOff's retrospective drug evaluation excluded chemical modifications, whereas Martinelli examined activity classification using both sequence and modification information. The tools differ in their inputs, prediction tasks, and training data. Their scores cannot simply be added together, nor can a high rank for an underlying sequence be assumed to persist after modification. The proposed strategy should therefore be compared with single-stage prediction and conventional screening using the same candidate set under standardized experimental conditions. Some candidates with intermediate or low ranks, or randomly selected candidates, should also be tested to determine whether sequences with high activity after modification have been missed. The effects of modification positions and combinations on stability and RNA interference activity still require experimental confirmation [7].

For delivery, a GalNAc-based liver-targeting strategy similar to that used for inclisiran could be considered, but cellular uptake and silencing activity would still need to be verified after conjugation of the new candidates [1, 8].

To evaluate the actual contribution of AI to screening efficiency, the same PCSK9 siRNA candidate set could be screened using AI-based ranking, conventional rule-based ranking, and random selection, with equal numbers of candidates from each group tested experimentally. Chemical modification patterns, GalNAc conjugation methods, test concentrations, and measurement times should be standardized across groups. The criterion for an effective candidate should be specified in advance, such as at least 70% PCSK9 mRNA knockdown. The proportion of effective candidates could first be compared across groups. Screening could then continue in predefined batches, recording the number of experimental tests and total time required to obtain a specified number of effective candidates. Efficiency assessments should also account for the data and time needed for model training or fine-tuning.

Modification optimization and validation in cells and animals could follow. PCSK9 mRNA knockdown can serve as an early screening measure, but a reduction in mRNA alone does not establish downstream functional improvement. Cell experiments should therefore also measure PCSK9 protein levels, cell-surface LDL receptor levels, and LDL uptake to assess whether silencing produces functional changes. Before animal validation, the candidate siRNA target sequence should be checked for compatibility with the PCSK9 transcript of the intended species, and candidates and animal models should be selected accordingly. An siRNA targeting human PCSK9 cannot be assumed to work directly in conventional mice. Subsequent evaluations should measure circulating PCSK9 protein, changes in LDL-C, duration of action, and safety [11]. Only after the necessary preclinical assessments have been completed and the requirements for clinical studies have been met should an ORION-like trial design be considered to compare a candidate drug with placebo as an addition to existing treatment in an appropriate patient population [3].

This approach benefits from an established mechanism of action for PCSK9 and a clinical basis for siRNA therapy. Studies of GalNAc-mediated liver-targeted delivery and chemical modifications also provide relevant technical foundations [1, 3, 7, 8]. Future studies can therefore assess the value of AI-assisted screening using specific measures such as candidate hit rate and experimental resources required.

A major challenge is model generalizability to new targets and modification designs. Martinelli's activity prediction study used 907 siRNAs for modeling and an external validation set of 57 siRNAs. Model performance varied across the 60%, 70%, and 80% activity classification thresholds. This variation means that task definitions must be considered when comparing studies [4]. AttSiOff initially performed poorly on the new target AGT and improved only after fine-tuning with an additional 600 AGT siRNAs. For PCSK9 screening, the model may need additional experimental data to perform reliably [5]. Martinelli's systematic review also notes that siRNA prediction studies differ in their datasets and evaluation methods [12]. AttSiOff would need to be evaluated on PCSK9 candidates. Predictions for chemically modified candidates would require a model that includes the relevant modification information, followed by experimental validation. Any time and resources spent collecting data, adapting the model, and testing it on independent data should count toward the total screening effort.

Silencing activity also provides an incomplete picture of drug performance: candidates must have suitable stability, effective delivery, and an acceptable safety profile. The predicted number of off-target binding sites cannot replace experimental safety evaluation [5, 7, 8]. Future work could compare screening approaches under controlled modification and delivery conditions, then use the resulting experimental data to update the models and assess whether candidate prioritization continues to improve.

4. Conclusion

Taken together, activity prediction studies, clinical evidence for PCSK9-targeted therapy, and research on modifications and delivery indicate a clear role for AI in early screening of lipid-lowering siRNAs: helping researchers prioritize candidates for experimental testing. However, current evidence does not establish that this workflow improves screening efficiency or therapeutic efficacy. Whether a candidate can become a drug also depends on its stability after modification, delivery, and safety, computational scores cannot replace these assessments.

This review is based mainly on representative studies and includes neither model replication nor experimental validation. Differences in experimental conditions across studies also require closer examination. Future studies could prospectively compare AI-assisted and conventional screening under defined modification and delivery conditions, using the proportion of effective candidates, experimental resources required, and duration of action to assess practical benefits. Newly generated experimental data could then be used to improve the models. Prioritized candidates would also require further validation of PCSK9 silencing, lipid-lowering effects, and safety to establish the practical value of computational screening.

References

[1] Marrs, J. C., & Anderson, S. L. (2024). Inclisiran for the treatment of hypercholesterolaemia. *Drugs in Context, 13*, Article 2023-12-3. https://doi.org/10.7573/dic.2023-12-3

[2] Seidah, N. G., & Prat, A. (2022). The multifaceted biology of PCSK9. *Endocrine Reviews, 43*(3), 558–582. https://doi.org/10.1210/endrev/bnab035

[3] Ray, K. K., Wright, R. S., Kallend, D., Koenig, W., Leiter, L. A., Raal, F. J., Bisch, J. A., Richardson, T., Jaros, M., Wijngaard, P. L. J., & Kastelein, J. J. P. (2020). Two phase 3 trials of inclisiran in patients with elevated LDL cholesterol. *The New England Journal of Medicine, 382*(16), 1507–1519. https://doi.org/10.1056/NEJMoa1912387

[4] Martinelli, D. D. (2024). From sequences to therapeutics: Using machine learning to predict chemically modified siRNA activity. *Genomics, 116*(2), Article 110815. https://doi.org/10.1016/j.ygeno.2024.110815

[5] Liu, B., Yuan, Y., Pan, X., Shen, H.-B., & Jin, C. (2024). AttSiOff: A self-attention-based approach on siRNA design with inhibition and off-target effect prediction. *Med-X, 2*(1), Article 5. https://doi.org/10.1007/s44258-024-00019-1

[6] Dmitriev, N. A., Chernov, P. V., Gongadze, I. S., Kovchina, V. I., Ivanov, V. N., Gusev, A. E., Shilovskiy, I. P., Kofiadi, I. A., & Khaitov, M. R. (2026). Refined design and liquid-phase assembly of GalNAc-siRNA conjugates: Comparative efficiency validation in PCSK9 targeting. *Molecules, 31*(3), Article 476. https://doi.org/10.3390/molecules31030476

[7] Foster, D. J., Brown, C. R., Shaikh, S., Trapp, C., Schlegel, M. K., Qian, K., Sehgal, A., Rajeev, K. G., Jadhav, V., Manoharan, M., Kuchimanchi, S., Maier, M. A., & Milstein, S. (2018). Advanced siRNA designs further improve in vivo performance of GalNAc-siRNA conjugates. *Molecular Therapy, 26*(3), 708–717. https://doi.org/10.1016/j.ymthe.2017.12.021

[8] Nair, J. K., Willoughby, J. L. S., Chan, A., Charisse, K., Alam, M. R., Wang, Q., Hoekstra, M., Kandasamy, P., Kel'in, A. V., Milstein, S., Taneja, N., O'Shea, J., Shaikh, S., Zhang, L., van der Sluis, R. J., Jung, M. E., Akinc, A., Hutabarat, R., Kuchimanchi, S., Fitzgerald, K., Zimmermann, T., van Berkel, T. J., Maier, M. A., Rajeev, K. G., & Manoharan, M. (2014). Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. *Journal of the American Chemical Society, 136*(49), 16958–16961. https://doi.org/10.1021/ja505986a

[9] Brown, C. R., Gupta, S., Qin, J., Racie, T., He, G., Lentini, S., Malone, R., Yu, M., Matsuda, S., Shulga-Morskaya, S., Nair, A. V., Theile, C. S., Schmidt, K., Shahraz, A., Goel, V., Parmar, R. G., Zlatev, I., Schlegel, M. K., Nair, J. K., . . . Jadhav, V. (2020). Investigating the pharmacodynamic durability of GalNAc–siRNA conjugates. *Nucleic Acids Research, 48*(21), 11827–11844. https://doi.org/10.1093/nar/gkaa670

[10] Wang, X., Wang, X., Varma, R. K., Beauchamp, L., Magdaleno, S., & Sendera, T. J. (2009). Selection of hyperfunctional siRNAs with improved potency and specificity. *Nucleic Acids Research, 37*(22), Article e152. https://doi.org/10.1093/nar/gkp864

[11] Frank-Kamenetsky, M., Grefhorst, A., Anderson, N. N., Racie, T. S., Bramlage, B., Akinc, A., Butler, D., Charisse, K., Dorkin, R., Fan, Y., Gamba-Vitalo, C., Hadwiger, P., Jayaraman, M., John, M., Jayaprakash, K. N., Maier, M., Nechev, L., Rajeev, K. G., Read, T., . . . Fitzgerald, K. (2008). Therapeutic RNAi targeting PCSK9 acutely lowers plasma cholesterol in rodents and LDL cholesterol in nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America, 105*(33), 11915–11920. https://doi.org/10.1073/pnas.0805434105

[12] Martinelli, D. D. (2024). Machine learning for siRNA efficiency prediction: A systematic review. *Health Sciences Review, 11*, Article 100157. https://doi.org/10.1016/j.hsr.2024.100157