Check for updates

SOFTWARE TOOL ARTICLE

YODEL: Peak calling software for HITS-CLIP data [version 1; peer review: 2 approved with reservations]

Lance E. Palmer¹ , Mitchell J. Weiss², Vikram R. Paralkar³¹Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN, 38105, USA²Department of Hematology, St. Jude Children's Research Hospital, Memphis, TN, 38105, USA³Division of Hematology/Oncology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA

V1 First published: 18 Jul 2017, 6:1138
<https://doi.org/10.12688/f1000research.11861.1>
Latest published: 18 Jul 2017, 6:1138
<https://doi.org/10.12688/f1000research.11861.1>

Abstract

YODEL is a peak calling software for analyzing RNA sequencing data generated by High-Throughput Sequencing of RNA isolated by Crosslinking Immunoprecipitation (HITS-CLIP; also known as CLIP-SEQ), a method to identify RNA-protein interactions genome-wide. We designed YODEL to analyze HITS-CLIP experiments, in which Argonaute proteins are immunoprecipitated, followed by sequencing of the associated RNA in order to identify bound microRNAs and their mRNA targets. The HITS-CLIP sequenced reads are mapped to the genome, and then read peaks are visualized where clustered sets of reads map to the same region. Several peak calling algorithms have been developed to define the boundaries of these peaks. In contrast to other peak callers for HITS-CLIP data, such as Piranha, YODEL does not map the starts of reads to fixed interval bins, but instead uses a heuristic approach to iteratively find the tallest point within a set clustered reads and examine bases upstream and downstream of that point until a peak has been determined. This allows the peak boundary to be defined more precisely than coordinates that are multiples of the bin size. Per-sample peak counts are also generated by YODEL, which quickly enables downstream differential representation analysis. YODEL is available at <https://github.com/LancePalmerStJude/YODEL/>.

Keywords

HITS-CLIP, CLIP-SEQ, peak caller

Open Peer Review

Approval Status ? ?

	1	2
version 1
18 Jul 2017	view	view

1. **Michele Trabucchi** , University of Côte d'Azur, Nice, France
Silvia Bottini, University of Côte d'Azur, Nice, France
2. **Neelanjan Mukherjee**, University of Colorado Anschutz Medical Campus, Aurora, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Lance E. Palmer (lance.palmer@stjude.org)

Author roles: **Palmer LE:** Conceptualization, Formal Analysis, Software, Writing – Original Draft Preparation; **Weiss MJ:** Funding Acquisition, Project Administration, Writing – Review & Editing; **Paralkar VR:** Conceptualization, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: LEP was supported by the Cancer Center Support (CORE) Grant (CA021765). VRP was supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grant (K08 DK08DK102533). MJW was supported by the NIDDK grant (R01 DK092318). This work was also supported by ALSAC.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2017 Palmer LE *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Palmer LE, Weiss MJ and Paralkar VR. **YODEL: Peak calling software for HITS-CLIP data [version 1; peer review: 2 approved with reservations]** F1000Research 2017, 6:1138 <https://doi.org/10.12688/f1000research.11861.1>

First published: 18 Jul 2017, 6:1138 <https://doi.org/10.12688/f1000research.11861.1>

Introduction

A peak caller that could accurately define a single peak amongst several samples was required to analyze High-Throughput Sequencing of RNA isolated by Crosslinking Immunoprecipitation (HITS-CLIP) (Darnell, 2010) data from fetal liver red blood cell precursors of *miR-144/451*^{-/-} and wild-type mice (Paralkar *et al.*, 2014; Paralkar VR, Palmer LE, Xu P, Lechauve C, Zhao G, Yao Y, Luan J, Wu G, Vourekas A, Mourelatos Z, Scheutz JD and Weiss MJ; unpublished study). Piranha (Uren *et al.*, 2012) is one such software that is commonly used to identify peaks generated by HITS-CLIP. However, Piranha bins the starts of reads and does not fully define a peak. Consequently, a large bin size may result in multiple peaks being combined into one peak. We found that the identification and resolution of peaks using Piranha was highly dependent on the background threshold (-a) and binSize (-b) parameters, and it was unclear how these parameters should be set in order to obtain the most biologically relevant information. We also found that running Piranha on multiple samples separately results in peak boundaries that may be quite different from sample to sample. Generating initial peak calls from a combined sample dataset creates a single standard set of peak boundaries for all samples, which simplifies downstream analysis. We therefore developed a peak calling algorithm, named YODEL, with the following properties: 1) Incorporate strand specificity (Piranha does this, but many other CHIP-SEQ peak callers do not); 2) Generate per-sample read counts for each peak; 3) Have parameters that have easily understandable implications when changed.

Methods

Input

The main input for the peak caller is a BED file generated by clusterBed from the BEDtools suite (Quinlan & Hall, 2010) with the -s option used (see [Supplementary material: Input file formats](#)). If multiple samples are to be analyzed simultaneously, the name field must contain the sample name or ID before the first colon, followed by the read ID or other descriptive text. In addition, a sample list must be designated (with the YODEL parameter -sampleList) (see [Supplementary material: Input file formats](#)). The sample list will identify which samples are to be included for peak calling. After peak calling, read counts for each peak in all samples will be calculated. If no sample list is provided, all reads will be treated as one sample. As an example of how to process HITS-CLIP FASTQ files to generate the input clustered BED file, see [Supplementary material: Analysis of HITS-CLIP data from Chi *et al.*, 2009](#).

Implementation

YODEL was written in Python and tested with Python version 2.7. YODEL processes each read cluster as it is encountered within the input clustered BED file. For each cluster, the base coverage at each position for all samples under examination for peak calling is determined. Position specific counts for all individual samples are calculated as well. Once the counts are tallied, the program iteratively identifies peaks until no additional peaks are found. The program identifies the position with the highest read count. If the read count is less than the minimum peak height (mph) than no further peak calling for the cluster is performed. From the position with the highest read count, bases upstream are analyzed one at a time. The lowest read count (lowestPoint) up to the

current base is tracked. Where dipTolerance (dt) and peakDipBuffer (pdb) are input parameters, if at any position the count is 0 or the count $\geq (\text{lowestPoint} + \text{peakDipBuffer}) * \text{dipTolerance}$, then the peak start has been determined and is recorded as the base position where the count was 0 or the base position of lowestPoint. This is repeated for bases downstream of the highest point to find the peak end. The peak summit is defined as the median of all the positions with the highest count. Two sets of peak boundaries (25% and 50%) are defined as the positions where the coverage is 25% and 50% of the highest point, and the maximum peak heights per sample are determined. The numbers of reads per peak are calculated (at both the 25% and 50% range) by determining the number of reads that overlap the peak by at least the input parameter binSize (bs). Typically, we have used the 25% peak boundary for downstream calculations. The peak counts are determined on a per sample basis, and for all the samples used in peak calling combined. Once a peak has been determined, the base coverage for positions within the full peak is set to 0 so that no further peaks are called overlapping it. The peak finding process (starting with finding the position with highest read count) is repeated until no more peaks are identified. After peaks are identified by YODEL, several filtering steps can be applied to remove low quality peaks (see [Supplementary material: Peak filtering, Supplementary Figure 2–Supplementary Figure 6](#)).

Operation

YODEL has been tested on both Windows and Linux running Python 2.7 with standard libraries. Some of the tools (e.g. Bedtools) used to generate input files are Linux or OS X specific. There is no minimum memory requirement for YODEL, but the size of any BED files sorted with the Linux sort command may be dependent on system memory. See '[Supplementary material: Analysis of HITS-CLIP data from Chi *et al.*, 2009](#)' for instructions on how to preprocess data and run YODEL.

Use case

The output of YODEL is described in the [Table 1](#). [Figure 1](#) shows a comparison of YODEL and Piranha output from HITS-CLIP

Table 1. Descriptions of output files.

File	Description
<i>prefix.bins.cov.full.bed</i>	Contains start and end coordinates for the entire peak, with the thickStart and thickEnd fields containing the 25% start and end coordinates, respectively
<i>prefix.bins.cov.bed</i>	Contains the 25% peak coordinates in the start and end fields and the 50% coordinates in thickStart and thickEnd fields.
<i>prefix.binCountsRC25.txt</i>	Read counts per peak defined by 25% peak
<i>prefix.binCountsRC50.txt</i>	Read counts per peak defined by 50% peak
<i>prefix.binCountsPH.txt</i>	Read counts covering position of maximum peak height for each individual sample

Figure 1. YODEL and Piranha peak calling comparisons in the *Cab39* and *Ywhaz* genes. IGV browser (Thorvaldsdóttir *et al.*, 2013) images showing YODEL output for HITS-CLIP data analyzing the 3' untranslated regions of *Cab39* (A) and *Ywhaz* (B) mRNAs in wild-type and *miR-451a/miR-144*^{-/-} mouse fetal livers at embryonic day 14.5. The coverage tracks (WT in blue, KO in red, and combined coverage in magenta) show combined sequencing reads from three animals of each genotype mapped to the mouse mm10 genome. The seedMatches track shows microRNA seed matches for *miR-451a*, *miR-144-3p* and the three most abundant erythroid microRNAs besides *miR-451a* (*miR-16-5p*, *miR-486a-5p* and *miR-122-5p*) (Paralkar VR, Palmer LE, Xu P, Lechauve C, Zhao G, Yao Y, Luan J, Wu G, Vourekas A, Mourelatos Z, Scheutz JD and Weiss MJ; unpublished results). Average peak counts per sample are shown for wild-type and *miR-451a/miR-144*^{-/-} erythroblasts using the YODEL more sensitive parameters (see below). The lower panels show YODEL (blue) and Piranha (green) peak boundaries with the indicated parameter settings. For YODEL, full boundaries are shown by thin lines and 25% boundaries by thick lines. YODEL less sensitive parameter settings: dipTolerance=2, peakDipBuffer=2; more sensitive settings: dipTolerance=1.5, peakDipBuffer=1; Both parameter settings: binSize=16, minPeakHeight=5. Note that the ability of Piranha to resolve different peaks representing distinct microRNA binding sites is highly dependent on parameter settings. Piranha parameters: a=background, threshold b=binSize. The default for Piranha is a=0.99 and b=32.

analysis of wild type and *miR144/451*^{-/-} fetal liver erythroblasts. Results from two different YODEL parameter settings are shown in blue in the lower half of the figure. *Cab39* (Figure 1A) (Godlewski *et al.*, 2010) and *Ywhaz* (Figure 1B) (Yu *et al.*, 2010) are two known *miR-451a* target mRNAs. For *Cab39*, the largest peak contains a *miR-451a* seed match that is not present in the knockout sample. Because Piranha bins the start of reads, the peak defined by Piranha may not actually include the seed match location, as is observed with the *Cab39* seed match. The failure of a peak to cover a seed match may prevent a microRNA from being assigned to a peak and subsequently interfere with downstream analysis. Also, the peak calling of Piranha is greatly influenced by bin size. A bin size of 32 (default parameter) is not able to resolve many individual peaks. In *Ywhaz* mRNA (panel B) YODEL detects three peaks around the *miR-451a* seed match. One Piranha setting (a=0.98, b=16) identified the three individual peaks, but

compromised detection of smaller peaks upstream of the predicted *miR-451a* binding site. Therefore for our *miR144/451*^{-/-} data set, YODEL was superior to Piranha in defining HITS-CLIP peaks.

We have also tested the YODEL software on a publically available HITS-CLIP data set. HITS-CLIP reads from mouse neocortex Argonaute immunoprecipitations (Chi *et al.* 2009) was retrieved from <http://ago.rockefeller.edu/rawdata.php>. The reads were pre-processed and run through YODEL, as described in the Supplementary material: Analysis of HITS-CLIP data from Chi *et al.*, 2009. We examined the first four genes identified by Table 1 in Chi *et al.* as potential target of microRNAs. Figure 2 shows a comparison of peak calling in the 3' UTR of the *Plod3* gene. Again, it is seen that binning starts of reads will cause potential microRNA seeds to be missed. Figure 3 shows the 3' UTR of the *Cd164* gene. See Supplementary Figure 1 for browser images of *Ctdsp1* and *Itgb1*.

Figure 2. YODEL and Piranha peak calling comparisons in the *Plod3* gene. IGV browser image of a portion of the *Plod3* 3' UTR. Mouse neocortex HITS-CLIP read coverage is shown along with peak calls from YODEL (lower panel blue bars) and Piranha (lower panel green bars). Relevant microRNA seeds (for seeds mapped see [Supplementary material](#): Seeds mapped for brain neocortex data) are shown in black. For YODEL, full boundaries are shown by thin lines and 25% boundaries by thick lines. YODEL parameters: dipTolerance=1.5, peakDipBuffer=1, binSize=16, minPeakHeight=5. Piranha parameters: a=background, threshold b=binSize. The default for Piranha is a=0.99 and b=32.

Figure 3. YODEL and Piranha peak calling comparisons in the *Cd164* gene. IGV browser image of a portion of the *Cd164* 3' UTR. Mouse neocortex HITS-CLIP read coverage is shown along with peak calls from YODEL (lower panel blue bars) and Piranha (lower panel green bars). Relevant microRNA seeds (For seeds mapped see [Supplementary material](#): Seeds mapped for brain neocortex data) are shown in black. For YODEL, full boundaries are shown by thin lines and 25% boundaries by thick lines. YODEL parameters: dipTolerance=1.5, peakDipBuffer=1, binSize=16, minPeakHeight=5. Piranha parameters: a=background, threshold b=binSize. The default for Piranha is a=0.99 and b=32.

Conclusions

We have designed a new peak-caller, termed YODEL, for analysis of RNA-seq data generated by HITS-CLIP-type experiments. Advantages of YODEL compared to Piranha, a program commonly used for the same purpose, include standardization of peak calls for comparative analysis of multiple samples, improved resolution of peak boundaries, and more consistent overlap between peak calls and microRNA seed matches.

Software and data availability

YODEL is a Python script and is available at: <https://github.com/LancePalmerStJude/YODEL/>

Archived source code as at time of publication: <https://doi.org/10.5281/zenodo.820635> (Palmer, 2017).

License: GPLv3

Input BED files for peak calling with the *miR-144/451* KO HITS-CLIP can be found in the Cab39_Ywhaz.allSamples.fullCollapsed.clusters.bed file within the archived source code listed above. This file contains the clustered BED file used for YODEL input (only around Cab39 and Ywhaz).

Sample HITS-CLIP data from the 130kD band from mouse neocortex samples was downloaded from <http://ago.rockefeller.edu/rawdata.php> (130kD Brain A-E samples).

The pre-processing BASH pipeline used to generate a clustered BED files, as well as some accessory scripts, can be found in the Ch2009 directory within the archived source code link listed above.

Competing interests

No competing interests were disclosed.

Grant information

LEP was supported by the Cancer Center Support (CORE) Grant (CA021765). VRP was supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grant (K08 5K08DK102533). MJW was supported by the NIDDK grant (R01 DK092318). This work was also supported by ALSAC.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplementary material

Supplementary File 1: YODEL supplementary text. This file contains supplementary text for this study. Within this file are descriptions of input files, publically available tools used, a second test data set analyzed by YODEL, and a description of a set of filters that can be used after peak calling.

[Click here to access the data.](#)

References

- Chi SW, Zang JB, Mele A, *et al.*: **Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps.** *Nature*. 2009; **460**(7254): 479–86.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Darnell RB: **HITS-CLIP: panoramic views of protein-RNA regulation in living cells.** *Wiley Interdiscip Rev RNA*. 2010; **1**(2): 266–286.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Godlewski J, Nowicki MO, Bronisz A, *et al.*: **MicroRNA-451 regulates LKB1/AMPK signaling and allows adaptation to metabolic stress in glioma cells.** *Mol Cell*. 2010; **37**(5): 620–632.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Palmer L: **LancePalmerStJude/YODEL: Yodel Initial Release.** *Zenodo*. 2017.
[Data Source](#)
- Paralkar VR, Luan J, Sridhar S, *et al.*: **Argonaute HITS-CLIP Reveals Global miRNA-mRNA Networks in Erythropoiesis.** *Blood*. 2014; **124**(21): 446.
[Reference Source](#)
- Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics*. 2010; **26**(6): 841–842.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Thorvaldsdóttir H, Robinson JT, Mesirov JP: **Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.** *Brief Bioinform*. 2013; **14**(2): 178–192.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Uren PJ, Bahrami-Samani E, Burns SC, *et al.*: **Site identification in high-throughput RNA-protein interaction data.** *Bioinformatics*. 2012; **28**(23): 3013–3020.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Yu D, dos Santos CO, Zhao G, *et al.*: **miR-451 protects against erythroid oxidant stress by repressing 14-3-3zeta.** *Genes Dev*. 2010; **24**(15): 1620–1633.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 16 August 2017

<https://doi.org/10.5256/f1000research.12817.r24815>

© 2017 Mukherjee N. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Neelanjan Mukherjee

Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

The authors present an alternative peak caller for HITS-CLIP data. While the idea is interesting and the examples are compelling, there is not sufficient analysis presented to determine the utility of YODEL.

Major:

1. The bench-marking is insufficient to evaluate the difference between YODEL and PIRANHA. The primary figures only have single examples. There needs to be a transcriptome-wide analysis to evaluate the performance.
2. The analysis should include some type of specificity/sensitivity analysis. It would be instructive to design "true" positives and "false" positive. Generally the "TRUE" positives could be thought of as miRNAs that are expressed in that system vs those that are not. Additionally, one can design 'decoy' seeds that are di-nucleotide shuffled seeds if the expressed miRNAs (that don't match the expressed miRNA seeds) and evaluate the number of counts relative to the actual expressed miRNA seed.

In the case of the KO, one could examine if the peaks called from WT data that contain seed matches to the KO miRNAs change in coverage (WT vs KO), particularly relative to compared to the peaks that contain seed matches to the expressed non-KO miRNAs. Comparing YODEL and PIRANHA in this analysis would be quite instructive.

Minor:

1. In the intro the authors describe three properties of YODEL. The first was:

"Incorporate strand specificity (Piranha does this, but many other CHIP-SEQ peak callers do not)"

I think this should be removed. Any peak finder for RNA interactions needs to be strand-

specific. I do not know why CHIP-SEQ peak callers are even mentioned, unless the authors believe this could be beneficial for CHIP-seq data. If so, they would need to compare to common CHIP-seq peak finders, though that would be a distraction in my opinion.

Is the rationale for developing the new software tool clearly explained?

Partly

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

No

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

No

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 25 July 2017

<https://doi.org/10.5256/f1000research.12817.r24325>

© 2017 Trabucchi M et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Michele Trabucchi

C3M (Mediterranean Centre for Molecular Medicine), INSERM (French National Institute of Health and Medical Research), University of Côte d'Azur, Nice, France

Silvia Bottini

C3M (Mediterranean Centre for Molecular Medicine), INSERM (French National Institute of Health and Medical Research), University of Côte d'Azur, Nice, France

The authors developed a new tool called YODEL to identify peaks from Ago2 HITS-CLIP data using a novel approach based on the identification of the peak summit of reads cluster and estimation

of the size of the peaks based on read coverage. The work is sound and interesting, however we have some concerns about the benchmarking of this new tool. Our major questions mainly refer to Bottini et al. (2017)¹. That should be cited.

Major suggestions/concerns:

Benchmarking:

1. The authors showed just for few selected targets that YODEL identifies peaks that include miRNAs seed matches, whereas Piranha did not. This should be shown at the genome-wide scale.

Inclusion of seed match sequences do not assure *per se* a better performance. In fact, seed matches can be included just by chance due to an overestimation of the peaks size. To rule out the possibility, the authors should show the distribution of the peaks length by the two programs on both entire datasets and calculate the correlation between peak length and number of seed matches.

2. For Ago2 CLIP-seq peak calling programs it is expected that miRNA-binding sites and cross-linked-dependent mutations position at the peak centers. How does YODEL perform compare to Piranha?
3. The authors defined two thresholds to assess the peak boundaries, namely 25% and 50% of the coverage of the highest point: how these two threshold have been assessed? Since the peaks boundaries is a primary concern for the authors, it should be explained and supported by analysis how they chose these two percentages. Therefore, we recommend to benchmark the thresholds by adding intermediate percentages and calculate the sensitivity toward seed match identified at the genome-wide scale.

Other major points:

1. The introduction needs improvement: a brief overview of the software/pipelines available to perform CLIP-seq data analysis and cite some reviews that explain all the steps of the data analysis, including Bottini et al. (2017)² and Uhl et al. (2017)³.
2. It should be clearly stated whether YODEL is able to find peaks enriched when comparing two conditions (differential CLIP) and/or only one condition.
3. In the Methods section all the parameters should be clearly stated and explained in the main text and not in the Supplementary Information. Furthermore, it should be added which kind of input data (not only the file format) are needed to run YODEL (replicates, IgG, control/KO ...).
4. Finally, it should be made clear whether YODEL can be applied to analyze only Ago2 HITS-CLIP data or also other RNA-binding proteins and why.

Minor points:

1. Supplemental figure 2 is missing.
2. About the first sentence on page 2 "A peak caller that could accurately define...." is odd. A

peak caller tool is always needed to identify peaks from CLIP-seq experiments, and not just for the specific case mentioned by the authors.

3. The sentence on page 2 "Where dipTolerance (dt) and peakDipBuffer (pbd) are input parameters..." is clumsy.
4. Some language misspelling such as: publically -> publicly.

References

1. Bottini S, Hamouda-Tekaya N, Tanasa B, Zaragosi LE, et al.: From benchmarking HITS-CLIP peak detection programs to a new method for identification of miRNA-binding sites from Ago2-CLIP data. *Nucleic Acids Res.* 2017; **45** (9): e71 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Bottini S, Pratella D, Grandjean V, Repetto E, et al.: Recent computational developments on CLIP-seq data analysis and microRNA targeting implications. *Brief Bioinform.* 2017. [PubMed Abstract](#) | [Publisher Full Text](#)
3. Uhl M, Houwaart T, Corrado G, Wright PR, et al.: Computational analysis of CLIP-seq data. *Methods.* 2017; **118-119**: 60-72 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the rationale for developing the new software tool clearly explained?

Partly

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Partly

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Partly

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research