

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

|                            |  |
|----------------------------|--|
| <b>TITLE (PROVISIONAL)</b> | EHDViz: Clinical Dashboard Development Using Open Source Technologies  |
| <b>AUTHORS</b>             | Badgeley, Marcus; Shameer, Khader; Glicksberg, Benjamin; Tomlinson, Max; Levin, Matthew; McCormick, Patrick; Kasarskis, Andrew; Reich, David; Dudley, Joel |

### VERSION 1 - REVIEW

|                        |                                  |
|------------------------|----------------------------------|
| <b>REVIEWER</b>        | Dingcheng Li<br>Mayo Clinic, USA |
| <b>REVIEW RETURNED</b> | 20-Dec-2015                      |

|                         |   |
|-------------------------|---|
| <b>GENERAL COMMENTS</b> | <p>This paper proposes to develop clinical dashboards using open source technologies, aiming at health assessments with interactive real-time visualization of heterogeneous biomedical, clinical and patient generated health data streams.</p> <p>This is really a significant idea in developing such clinical dashboards. Especially, it is released as an open-source visualization framework and we believe that the tool, named as EHDViz will become an indispensable toolkit for clinical data visualization.</p> <p>This work addresses the problem of the intergration of diverse data elements and provide a common platform for visulization. In addition, it looks that EHDViz is not only based on vital signs and status, but also has incorporated predictive laboratory values, physician orders and medications. So, EHDViz provides an EHR-agnostic visualization and does real-time assistance in identifying patients with a rich and visual aid.</p> <p>It provides three experimental results in details. They clearly show that EHDViz enjoys hight quality, high-velocity data visualization in a scalable manner.</p> <p>But I have a few criticisms on the construction of this paper.</p> <ol style="list-style-type: none"> <li>1. It looks not so clear how this tool is developed and what the exact methodologies are. It only says that they make use of open-source technologies, such as R packages. But the authors didn't describe the details in their methods section. Compared with other sections, Methods only invovles two short paragraphs.</li> <li>2. When related works are described, only the names of those tools are listed and we are not clear what are the essential differences between other workds and this work. They only say that EHDViz perform much better. But no specific details about why EHDViz can do a better job than others.</li> </ol> |
|-------------------------|---|

|  |   |
|--|---|
|  | 3. Some descriptions are a little bit hard to understand. Maybe, it is wise to write simpler sentences. |
|--|---|

|                        |  |
|------------------------|--|
| <b>REVIEWER</b>        | Enno T. van der Velde, PhD<br>Department of Cardiology, Leiden University Medical Center,<br>Leiden, The Netherlands |
| <b>REVIEW RETURNED</b> | 21-Dec-2015  |

|                         |   |
|-------------------------|---|
| <b>GENERAL COMMENTS</b> | <p>The paper describes a problem that has arisen in recent years, namely that the doctor/caretaker needs to be able to view and judge healthcare data from many different sources (EHR, patient generated data, lab data) quickly and sufficiently. The discussed Clinical Dashboard application, developed by the authors, seems to be a useful tool to solve this problem. However, I have the following comments:</p> <ul style="list-style-type: none"> <li>- the English text should be improved drastically. The text contains many, many typing errors, wrong grammar, incorrect wordings, and sentences that are difficult or impossible to understand the meaning.</li> <li>-- For instance in the subsection 'Visualization of risk estimation..': "Inpatient adverse events" (=???); "antecedents retrospectively" (= ???; maybe respectively?). What are "outpatient wellness visits" (Results section, first paragraph)?</li> <li>- The paper is more a demonstration project than a real implementation. The dashboard does not yet have a connection to a real electronic health record system, which would make it more interesting, but also more complicated with respect to taking patient privacy and security measures.</li> <li>- The Introduction section is very long, too long compared to for instance the Method section.</li> <li>- The Methods section is too short, and does not give enough information to allow the study to be repeated. Which fitness monitoring devices (plural!) were used? What is an internal API?</li> <li>- Some abbreviations are not explained (AUROC, and other).</li> <li>- Some url's in the text are not functional (e.g.: the 2 url's at the of the Results/Dashboard 3 section).</li> </ul> |
|-------------------------|---|

### VERSION 1 – AUTHOR RESPONSE

#### Reviewer: 1

Reviewer Name

Dingcheng Li

Institution and Country

Mayo Clinic, USA

Please state any competing interests or state 'None declared':

None

Please leave your comments for the authors below

This paper proposes to develop clinical dashboards using open source technologies, aiming at health assessments with interactive real-time visualization of heterogeneous biomedical, clinical and patient generated health data streams. This is a significant idea in developing such clinical dashboards.

Especially, it is released as an open-source visualization framework, and we believe that the tool, named as EHDViz will become an indispensable toolkit for clinical data visualization.

This work addresses the problem of the integration of diverse data elements and provides a common

platform for visualization. Also, it looks that EHDViz is not only based on vital signs and status but also has incorporated predictive laboratory values, physician orders and medications. So, EHDViz provides an EHR-agnostic visualization and does real-time assistance in identifying patients with a rich and visual aid. It provides three experimental results in details. They clearly show that EHDViz enjoys high quality, high-velocity data visualization in a scalable manner.

*We would like to thank the reviewer for the encouraging comments.*

But I have a few criticisms on the construction of this paper.

1. It looks not so clear how this tool is developed and what the exact methodologies are. It only says that they make use of open-source technologies, such as R packages. But the authors didn't describe the details in their methods section. Compared with other sections, Methods only involves two short paragraphs.

*We have now expanded the "Methods" section with sufficient details, and technical concepts are explained in simple language. In the revised the version, we have now introduced the following new sections under Methods:*

- *Description of EHDViz Framework*
- *Data handling in EHDViz*
- *Input and output specifications of EHDViz*

2. When related works are described, only the names of those tools are listed, and we are not clear what are the essential differences between other works and this work. They only say that EHDViz performs much better. But no specific details about why EHDViz can do a better job than others.

*We agree with the reviewer that a comparative analysis is an important detail; we inadvertently missed some of the tool with similar capabilities in the manuscript. A new section on the comparison with different tools (EventFlow, Lifelines, Lifelines2 and Harvest (See Refs 1-4)) for real-time clinical data visualization are added in the revision. We compiled EHDviz using various factors including deployment, technology stack, extensibility options, options for real-time data display, options for incorporating risk estimations using predictive analytics and machine learning. To the best of our understanding, EHDViz has better extensibility as it can handle multiple data types and can use any R package for real-time visualization and can be integrated with any existing EHR systems in a vendor agnostic approach. While an objective comparison will provide more nuanced details including the feature like usability, such a comparison would need additional resources and could be performed like a future trial for evaluation of HealthIT applications for clinical data visualization.*

*Comparison of various features of real-time clinical data visualization tools are now summarized in a new section under the Discussion of revised version:*

- *Comparison with related healthcare data visualization applications:*

3. Some descriptions are a little bit hard to understand. Maybe, it is wise to write simpler sentences.

*We have made several edits to improve the readability and revised subheadings of sections to provide more clarity and corrected grammatical and phrasing errors. A total of 11 new references is also included. Figure 1 is now replaced with a new flow-chart which is more descriptive compared to the infographic based Figure in the last version. The full panel Figure is now replaced as independent figures to show the details/depth of visualization capabilities of EHDViz. We hope these revisions further simplify and explain the concept, utilities, details/depth of graphical features of EHDViz.*

## Reviewer: 2

### Reviewer Name

Enno T. van der Velde, PhD

### Institution and Country

Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands

Please state any competing interests or state 'None declared':

None declared

Please leave your comments for the authors below

The paper describes a problem that has arisen in recent years, namely that the doctor/caretaker needs to be able to view and judge healthcare data from many different sources (EHR, patient generated data, lab data) quickly and sufficiently. The discussed Clinical Dashboard application, developed by the authors, seems to be a useful tool to solve this problem. However, I have the following comments:

*We would like to thank the reviewer for encouraging comments.*

- The English text should be improved drastically. The text contains many, many typing errors, wrong grammar, incorrect wordings, and sentences that are difficult or impossible to understand the meaning.

*We have made several edits to improve the readability and revised subheadings of sections to provide more clarity and corrected grammatical and phrasing errors.*

-- For instance in the subsection 'Visualization of risk estimation..': "Inpatient adverse events" (=???)  
"antecedents retrospectively" (= ???; maybe respectively?). What are "outpatient wellness visits" (Results section, first paragraph)?

*These sections are revised.*

- The paper is more a demonstration project than a real implementation. The dashboard does not yet have a connection to a real electronic health record system, which would make it more interesting, but also more complicated on taking patient privacy and security measures.

*We would like to thank the reviewer for encouraging critique of our work. The examples we have shown in the manuscript are of demonstration in nature. The idea is to prime the physician and biomedical informatics communities to embrace options in the open source technology world and adopt and adapt it for developing high-quality, high-latency clinical applications. As we are now entering an era of EHR-centered care, development of customized clinical dashboard that layer on top of EHR is of critical importance for innovation, rapid prototyping of clinical software applications, and we hope EHDViz could be one of the first examples of such capabilities using diverse data feeds. With this manuscript, we introduce EHDViz as a cost-effective alternative to other clinical dashboard development solutions. Further, EHDViz, as an open platform will be useful as a tool for teaching computational medicine and encourage medical students to create "mock-dashboards" and use as an effective teaching aid by providing the source code and sample dashboards in public domain.*

- The Introduction section is very long, too long compared to for instance the Method section.

*We have now expanded the methods section with sufficient details, and additional references are added.*

- The Methods section is too short and does not give enough information to allow the study to be

repeated. Which fitness monitoring devices (plural!) were used? What is an internal API?

*We have now expanded the methods section with sufficient details, and technical concepts are expressed in a simpler language.*

- Some abbreviations are not explained (AUROC, and other).

*We have now expanded all abbreviations in the manuscript.*

- Some URL's in the text are not functional (e.g.: the two url at the of the Results/Dashboard 3 section).

*We have now fixed all URLs in the manuscript.*

**References:**

1 Monroe, M., Rongjian, L., Hanseung, L., Plaisant, C. & Shneiderman, B. Temporal Event Sequence Simplification. Visualization and Computer Graphics, IEEE Transactions on 19, 2227-2236, doi:10.1109/TVCG.2013.200 (2013).  
 2 Pennington, J. W. et al. Harvest: an open platform for developing web-based biomedical data discovery and reporting applications. Journal of the American Medical Informatics Association: JAMIA 21, 379-383, doi:10.1136/amiainl-2013-001825 (2014).  
 3 Plaisant, C. et al. LifeLines: using visualization to enhance navigation and analysis of patient records. Proc AMIA Symp, 76-80 (1998).  
 4 Wang, T. D. et al. Temporal summaries: Supporting temporal categorical searching, aggregation, and comparison. IEEE Trans Vis Comput Graph 15, 1049-1056, doi:10.1109/TVCG.2009.187 (2009).

**VERSION 2 – REVIEW**

|                        |   |
|------------------------|---|
| <b>REVIEWER</b>        | Dingcheng Li<br>Division of Biomedical Statistics and Informatics, Mayo Clinic, USA |
| <b>REVIEW RETURNED</b> | 12-Feb-2016   |

|                         |   |
|-------------------------|---|
| <b>GENERAL COMMENTS</b> | From this version, I can see that the authors have made thorough revisions and improvements. According to my perspective, the authors have addressed all comments, including adding more references. One point I want to know is where the response letters to each reviewer's comments? Did I miss it? Although I can see the old version, the new version and the revised parts, it would be better for reviewers to see how the authors respond to each comment. |
|-------------------------|---|

|                        |  |
|------------------------|--|
| <b>REVIEWER</b>        | Enno T. van der Velde, PhD<br>Leiden University Medical Center, HeartLungCenter<br>Leiden, The Netherlands |
| <b>REVIEW RETURNED</b> | 21-Feb-2016  |

|                         |  |
|-------------------------|--|
| <b>GENERAL COMMENTS</b> | Manuscript is greatly improved, with for instance extensive and clear methods section.<br>Just a few remarks now:<br>In some sentences of the manuscript the grammar could be improved, e.g.:<br>p 5, line 47: comma after devices<br>p 9, line 27: empowers -> empower<br>p9, lines 54+56: what is data cleansing?<br>p 10, line 9: "A custom, the plot-stitching algorithm, ...: grammar |
|-------------------------|--|

|  |   |
|--|---|
|  | <p>needs to be improved</p> <p>p 10, line 34: ".. can be indexed various clinical ..." improve grammar</p> <p>p 10, line 54: ".. clinical operations "= ??</p> <p>p 11, line 3: ".. patients cohorts TO match .." (the word 'to' is missing)</p> <p>p 11, line 58: sentence starting with "1) data from a single patient.." improve grammar</p> <p>p 12, line 11: ".. API capable of ..": capabel of what?</p> <p>Etcetera.. So please review the whole manuscript to improve grammar.</p> <p>Secondly: some of the references are incomplete, with missing Journal name etcetera. For instance ref 18. Please check all references for completeness.</p> |
|--|---|

## VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name

Dingcheng Li

Institution and Country

Mayo Clinic, USA

Please state any competing interests or state 'None declared':

None

Please leave your comments for the authors below

This paper proposes to develop clinical dashboards using open source technologies, aiming at health assessments with interactive real-time visualization of heterogeneous biomedical, clinical and patient generated health data streams. This is really a significant idea in developing such clinical dashboards. Especially, it is released as an open-source visualization framework and we believe that the tool, named as EHDViz will become an indispensable toolkit for clinical data visualization.

This work addresses the problem of the integration of diverse data elements and provide a common platform for visualization. In addition, it looks that EHDViz is not only based on vital signs and status, but also has incorporated predictive laboratory values, physician orders and medications. So, EHDViz provides an EHR-agnostic visualization and does real-time assistance in identifying patients with a rich and visual aid. It provides three experimental results in details. They clearly show that EHDViz enjoys high quality, high-velocity data visualization in a scalable manner.

We would like to thank the reviewer for the encouraging comments.

But I have a few criticisms on the construction of this paper.

1. It looks not so clear how this tool is developed and what the exact methodologies are. It only says that they make use of open-source technologies, such as R packages. But the authors didn't describe the details in their methods section. Compared with other sections, Methods only involves two short paragraphs.

We have now expanded the "Methods" section with sufficient details, and technical concepts are explained in simple language. In the revised the version, we have now introduced the following new sections under Methods:

- Description of EHDViz Framework
- Data handling in EHDViz
- Input and output specification s of EHDViz

2. When related works are described, only the names of those tools are listed and we are not clear

what are the essential differences between other works and this work. They only say that EHDViz perform much better. But no specific details about why EHDViz can do a better job than others.

We agree with the reviewer that a comparative analysis is an important detail; we inadvertently missed some of the tool with similar capabilities in the manuscript. A new section on the comparison with different tools for real-time clinical data visualization is added in the revision. We compiled EHDViz using various factors including deployment, technology stack, extensibility options, options for real-time data display, options for incorporating risk estimations using predictive analytics and machine learning (Table R1). To the best of our understanding, EHDViz has better extensibility as it can handle multiple data types and can use any R package for real-time visualization and can be integrated with any existing EHR systems in a vendor agnostic approach. While, an objective comparison will provide more nuanced details including the feature like usability, such a comparison would need additional resources and could be performed like a future trial for evaluation of HealthIT applications for clinical data visualization.

| Tool                    | Deployment | Technology   | Extensibility | Real-time display | Machine learning |
|-------------------------|------------|--------------|---------------|-------------------|------------------|
| EHDViz                  | Web-based  | R and Python | Yes           | Yes               | Yes              |
| EventFlow <sup>1</sup>  | Desktop    | Java         | Limited       | Yes               | No               |
| HARVEST <sup>2</sup>    | Web-based  | Python       | Yes           | Limited           | No               |
| LifeLines <sup>3</sup>  | Desktop    | Java         | Limited       | Yes               | No               |
| LifeLines <sup>24</sup> | Desktop    | Java         | Limited       | Yes               | No               |

Table R1: Comparison of various features of real-time clinical data visualization tools

These results are now summarized under a new section under the Discussion of revised version:  
 • Comparison with related healthcare data visualization applications:

3. Some descriptions are a little bit hard to understand. Maybe, it is wise to write simpler sentences.

We have made several edits to improve the readability and revised subheadings of sections to provide more clarity and corrected grammatical and phrasing errors. A total of 11 new references are also included. Figure 1 is now replaced with a new flow-chart which is more descriptive compared to the infographics based figure in the last version. The full panel figure is now replaced as independent figures to show the details/depth of visualization capabilities of EHDViz. We hope these revision further simplify and explain the concept, utilities, details/depth of graphical features of EHDViz.

Reviewer: 2

Reviewer Name

Enno T. van der Velde, PhD

Institution and Country

Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands

Please state any competing interests or state 'None declared':

None declared

Please leave your comments for the authors below

The paper describes a problem that has arisen in recent years, namely that the doctor/caretaker needs to be able to view and judge healthcare data from many different sources (EHR, patient generated data, lab data) quickly and sufficiently. The discussed Clinical Dashboard application, developed by the authors, seems to be a useful tool to solve this problem. However, I have the following

comments:

We would like to thank the reviewer for encouraging comments.

- the English text should be improved drastically. The text contains many, many typing errors, wrong grammar, incorrect wordings, and sentences that are difficult or impossible to understand the meaning.

We have made several edits to improve the readability and revised subheadings of sections to provide more clarity and corrected grammatical and phrasing errors.

-- For instance in the subsection "Visualization of risk estimation..": "Inpatient adverse events" (=???); "antecedents retrospectively" (= ???; maybe respectively?). What are "outpatient wellness visits" (Results section, first paragraph)?

These sections are revised.

- The paper is more a demonstration project than a real implementation. The dashboard does not yet have a connection to a real electronic health record system, which would make it more interesting, but also more complicated with respect to taking patient privacy and security measures.

We would like to thank the reviewer for encouraging critique of our work. The examples we have shown in the manuscript are of demonstration in nature. The idea is to prime the physician and biomedical informatics communities to embrace options in the open source technology world and adopt and adapt it for developing high-quality, high-latency clinical applications. As we are now entering an era of EHR-centered care, development of customized clinical dashboard that layer on top of EHR is of critical importance for innovation, rapid prototyping of clinical software applications, and we hope EHDViz could be one of the first examples of such capabilities using diverse data feeds. With this manuscript, we introduce EHDViz as a cost-effective alternative to other clinical dashboard development solutions. Further, EHDViz, as an open platform will be useful as a tool for teaching computational medicine and encourage medical students to create "mock-dashboards" and use as an effective teaching aid.

- The Introduction section is very long, too long compared to for instance the Method section.

We have now expanded the methods section with sufficient details and additional references are added.

- The Methods section is too short, and does not give enough information to allow the study to be repeated. Which fitness monitoring devices (plural!) were used? What is an internal API?

We have now expanded the methods section with sufficient details and technical concepts are expressed in a simpler language.

- Some abbreviations are not explained (AUROC, and other).

We have now expanded all abbreviations in the manuscript.

- Some url's in the text are not functional (e.g.: the 2 url's at the of the Results/Dashboard 3 section.

We have now fixed all URLs in the manuscript.

References:

- 1 Monroe, M., Rongjian, L., Hanseung, L., Plaisant, C. & Shneiderman, B. Temporal Event Sequence Simplification. *Visualization and Computer Graphics, IEEE Transactions on* 19, 2227-2236, doi:10.1109/TVCG.2013.200 (2013).
- 2 Pennington, J. W. et al. Harvest: an open platform for developing web-based biomedical data discovery and reporting applications. *Journal of the American Medical Informatics Association : JAMIA* 21, 379-383, doi:10.1136/amiajnl-2013-001825 (2014).
- 3 Plaisant, C. et al. LifeLines: using visualization to enhance navigation and analysis of patient records. *Proc AMIA Symp*, 76-80 (1998).
- 4 Wang, T. D. et al. Temporal summaries: supporting temporal categorical searching, aggregation and comparison. *IEEE Trans Vis Comput Graph* 15, 1049-1056, doi:10.1109/TVCG.2009.187 (2009).