

# PROCEEDINGS OF SPIE

[SPIEDigitalLibrary.org/conference-proceedings-of-spie](https://SPIEDigitalLibrary.org/conference-proceedings-of-spie)

## Are fiducial registration error and target registration error correlated? SciKit-SurgeryFRED for teaching and research

Thompson, Stephen, Dowrick, Tom, Ahmad, Mian, Opie, Jeremy, Clarkson, Matthew

Stephen Thompson, Tom Dowrick, Mian Ahmad, Jeremy Opie, Matthew J. Clarkson, "Are fiducial registration error and target registration error correlated? SciKit-SurgeryFRED for teaching and research," Proc. SPIE 11598, Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling, 115980U (15 February 2021); doi: 10.1117/12.2580159

**SPIE.**

Event: SPIE Medical Imaging, 2021, Online Only

# Are fiducial registration error and target registration error correlated? SciKit-SurgeryFRED for teaching and research.

Stephen Thompson<sup>1</sup>, Tom Dowrick<sup>1</sup>, Mian Ahmad<sup>1</sup>, Jeremy Opie<sup>1</sup>, Matthew J. Clarkson<sup>1</sup>

<sup>1</sup>Wellcome/EPSRC Centre for Interventional and Surgical Science, University College London, United Kingdom

## ABSTRACT

Understanding the relationship between fiducial registration error (FRE) and target registration error (TRE) is important for the correct use of interventional guidance systems. Whilst it is well established that TRE is statistically independent of FRE, system users still struggle against the intuitive assumption that a low FRE indicates a low TRE. We present the SciKit-Surgery Fiducial Registration Educational Demonstrator and describe its use. SciKit-SurgeryFRED was developed to enable remote teaching of key concepts in image registration. SciKit-SurgeryFRED also supports research into user interface design for image registration systems.

SciKit-SurgeryFRED can be used to enable remote tutorials covering the statistics relevant to image guided interventions. Students are able to place fiducial markers on pre and intra-operative images and observe the effects of changes in marker geometry, marker count, and fiducial localisation error on TRE and FRE. SciKit-SurgeryFRED also calculates statistical measures for the expected values of TRE and FRE. Because many registrations can be performed quickly the students can then explore potential correlations between the different statistics.

SciKit-SurgeryFRED also implements a registration based game, where participants are rewarded for complete treatment of a clinical target, whilst minimising the treatment margin. We used this game to perform a remote study on registration and simulated ablation, measuring how user performance changes depending on what error statistics are made available. The results support the assumption that knowing the exact value of target registration error leads to better treatment. Display of other statistics did not have a significant impact on the treatment performance.

## 1. INTRODUCTION

The fact that fiducial registration error (FRE) is uncorrelated with target registration error (TRE) is well established.<sup>10</sup> In spite of this, many students and users of clinical guidance systems struggle to correctly interpret FRE and residual errors in general. SciKit-SurgeryFRED (Fiducial Registration Educational Demonstrator)<sup>20</sup> was developed using the SciKit-Surgery<sup>22</sup> libraries to probe the causes of misconceptions about FRE and TRE. SciKit-SurgeryFRED supports online learning and provides tools to enable research into user interface design for image guidance systems.

Faced with the need to deliver teaching and research remotely during 2020's SARS-CoV-2<sup>7</sup> travel restrictions, we developed a new online tutorial\* on fiducial based registration. The tutorial utilises SciKit-SurgeryFRED which is now implemented as a browser based web-app. SciKit-SurgeryFRED can be accessed on-line† or deployed locally. Rather than simply telling the students that FRE is uncorrelated with TRE we decided it would be more effective to use divergent questioning strategies<sup>25</sup> within the application, analysis, synthesis and evaluation levels of Bloom's taxonomy:<sup>1</sup> hence the question posed within the title of this paper.

We also took the opportunity to develop a serious game based on SciKit-SurgeryFRED. Serious games are games designed to be engaging to play, whilst setting out to achieve an objective beyond pure entertainment. In the case of SciKit-SurgeryFRED, the objective is to gather data on how clinical decision making might be influenced by the communication of registration errors. Serious games have significant potential as a training tool for clinicians.<sup>4,13</sup> Serious games have also been shown to be useful in assessing the usability of human computer

\* [https://mph0026.readthedocs.io/en/latest/summerschool/registration\\_demo.html](https://mph0026.readthedocs.io/en/latest/summerschool/registration_demo.html)

† <https://scikit-surgeryfred.ew.r.appspot.com/>

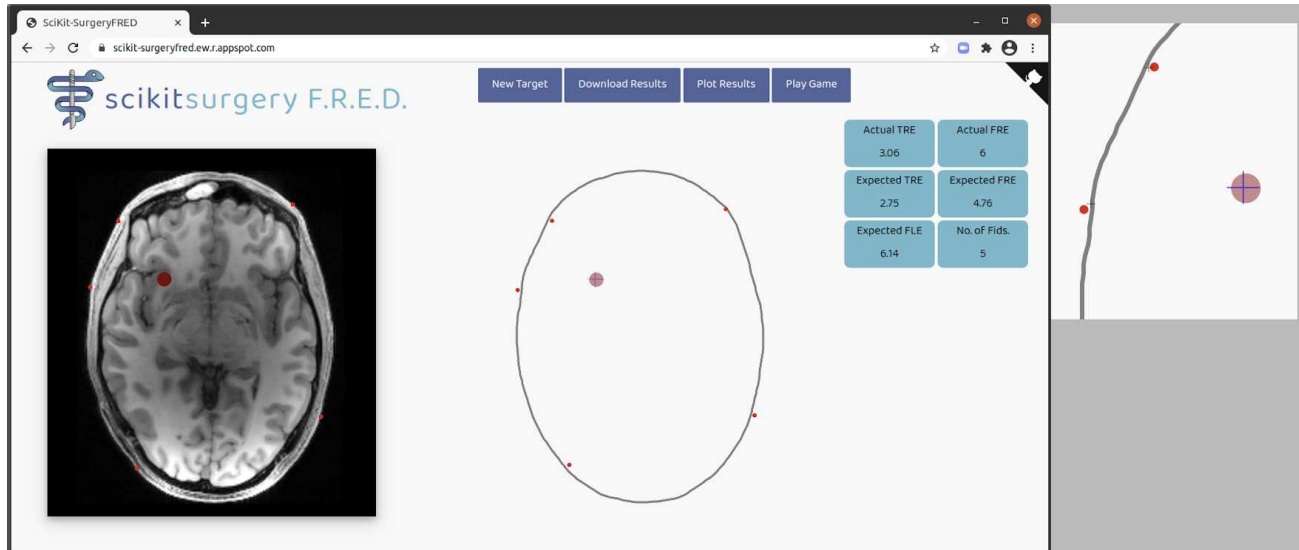


Figure 1: SciKit-SurgeryFRED graphical user interface after 6 fiducial markers placed. The red circle in the pre-operative image (left) represents a clinical target, which is located in the intra-operative space (middle) by fiducial based registration, using the fiducial markers (also red). FLE is added to each marker in the intra-operative image, this is more clearly visible on the zoomed in image at the right. The resulting registration results in a TRE, shown by the misalignment of the red circle and crosshair on the enlarged image at right. TRE and other statistics are shown in boxes to the right of the intra-operative image.

interfaces,<sup>6</sup> which is of particular relevance to computer aided surgery. The results of the game indicate that the choice of error statistic shown can affect clinical decision making.

In this paper we introduce the SciKit-SurgeryFRED application and show how it can be used as an educational tool and to perform studies to measure registration and ablation performance. By default SciKit-SurgeryFRED implements fiducial localisation error (FLE) as an isotropic, normally distributed, independent variable. In this paper we show how SciKit-SurgeryFRED can be modified to implement anisotropic and systematic errors, and present the results of these modifications. It is our plan to build on the software to perform larger studies; to investigate more realistic models of FLE; to investigate other registration approaches; and other ways of communicating registration error. SciKit-SurgeryFRED is entirely open source software and we encourage researchers and educators to use and contribute to it.

## 2. METHODS AND IMPLEMENTATION

SciKit-SurgeryFRED implements a minimal user interface, Figure 1, to demonstrate registration of a pre-operative image to intra-operative space. At startup a target is placed at a random location within the pre-operative image, shown as a red circle. The standard deviation of the FLE is randomly sampled from a uniform distribution between 0.5 and 5.0 pixels<sup>‡</sup>. By default the FLE is modelled as an isotropic (in 3 dimensions), normally distributed, and independent random variable, though this can be easily changed. The variance of the FLE is shown to the right of the intra-operative image (as expected value), along with the number of fiducial markers.

Clicking either image adds a fiducial marker to both images. By default the marker is added to the pre-operative image with no FLE. FLE is added to the marker location in the intra-operative image, visualised as the misalignment between the red circle centre and the cross-hair. Once sufficient markers are placed (> 2) the two sets of markers are registered using least squares fitting as described by Arun et al.<sup>2</sup> The expected values

<sup>‡</sup>All units are in pixels, as the concepts being explored do not depend on the units used. Figure 1 shows a brain MRI 458x512 pixels, by design SciKit-SurgeryFRED should work with other images of arbitrary dimensions.



Figure 2: Plots of TRE against various error measures, generated by 205 user registrations.

(variance) of the FRE and TRE are calculated using equations described by Fitzpatrick et al.<sup>9</sup> The student can use this interface together with an online tutorial to explore the relationships between the various statistics and error measures. The user can keep adding as many fiducial markers as they like for a given target. Pressing the “New Target” button will place a new target at a random position within the pre-operative image and randomly sample a new standard deviation for the FLE.

It is straightforward to explore how both TRE and FRE change in response to both the number and geometry of the fiducial markers as is well established in the literature.<sup>9,28</sup> It is also simple and instructive to create degenerate marker geometries, for example a linear arrangement can produce low FRE but extreme values of TRE.

During use the registration results are stored in the browser. At any time the user can press the “Plot Results” button to quickly create a set of plots showing the relationship between the actual TRE and the various statistics. An example plot is shown in Figure 2. The student is usually able to see first hand that, as expected, FRE is uncorrelated with TRE.

SciKit-SurgeryFRED does not implement tests of statistical significance to avoid excessive software dependencies, so it is often useful for the students to download the registration results and perform some statistical analysis. This is facilitated by the “Download Results” button which allows the user to get the results as a comma separated variable format. For example the results in Figure 2 may prompt some students to ask if there is in fact a correlation between FRE and TRE, as there is an apparent though slight increase in TRE with FRE. This can quickly be dispelled by asking the student to download the results and perform a test of significance on the data.

## 2.1 Game Based Registration Study

Once the students have an understanding of what statistics can be used to estimate TRE, we wanted to test how knowledge of a particular statistic affects optimal treatment planning. We designed a serious game to do this. During this game the students are asked to perform a registration with a maximum of 6 fiducial markers, then set a treatment margin and ablate the target. The goal is to treat 100% of the target with minimal ablation of surrounding tissue. A score of 1000 was awarded for complete ablation and 0 for anything less. From this starting

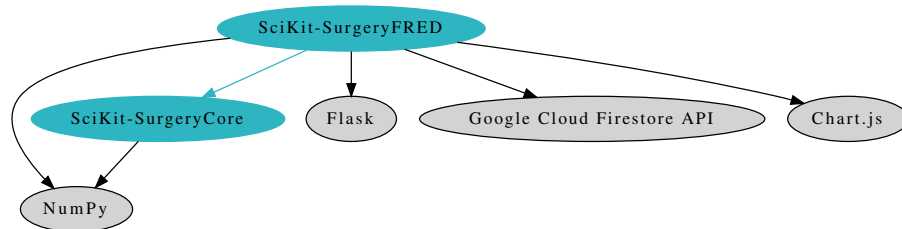


Figure 3: Software dependencies of SciKit-SurgeryFRED. The registration algorithms and statistical measures of TRE are imported from SciKit-SurgeryCore and are shared with clinical applications built on SciKit-Surgery. The user interface is a Flask based web application, using Chart.js for plotting functionality.

score 10 times the percentage volume of any surrounding tissues ablated is subtracted. Successful treatments (100% ablation) should therefore score between 0 and 1000, with a larger margin giving a lower score. The optimum (minimum successful) treatment margin is the actual TRE for a given registration. Any treatment that fails to treat 100% of the target will receive a negative score.

Each student participated in the game, performing 20 simulated ablations. For the first four ablations they were told the actual TRE for training and to supply baseline data. After that they performed 16 more ablations and were shown one of four randomly selected statistics on which to base their decision, the expected values of the TRE/FRE, the actual FRE, the expected value of the FLE. These statistics were chosen as these are either known or can be estimated for a clinical procedure. Each statistic was shown 4 times, though the order was randomised for each participant. The scores for each ablation were recorded, yielding 20 data points for each participant.

## 2.2 Software Implementation

SciKit-SurgeryFRED is part of the SciKit-Surgery<sup>22</sup> family of libraries. In common with SciKit-Surgery the majority of SciKit-SurgeryFRED is implemented in Python. Python was chosen as it combines sufficient features for clinical applications such as the SmartLiver system,<sup>17</sup> whilst remaining easy enough for students to learn and contribute to. In contrast, although platforms built using C++ provide power and flexibility, the choice of language creates a barrier to learning the key concepts of image guided surgery.<sup>19</sup> A key design goal of SciKit-Surgery is to keep individual libraries compact and orthogonal,<sup>12</sup> simplifying dependency structures. Based on analysis using cloc<sup>§</sup> SciKit-SurgeryFRED consists of 1479 lines of Python code. The user interface is implemented in HTML5 and JavaScript, enabling multiple simple deployment options. Again, using cloc, there are 957 lines of HTML5 and JavaScript. These numbers are similar to the other SciKit-Surgery libraries which typically have around 2000 lines of code.<sup>22</sup>

Figure 3 shows the direct dependencies of SciKit-SurgeryFRED. The key functional dependency is SciKit-SurgeryCore.<sup>5</sup> SciKit-SurgeryCore implements matched point based registration<sup>2</sup> together with the calculation of expected FLE and TRE (equations 10 and 31 from Fitzpatrick et al.<sup>9</sup>). NumPy<sup>11</sup> is used for array handling. Flask<sup>¶</sup> provides the web application framework to enable the browser based user interface to communicate with the Python based back end. The user interface communicates with the back end with a series of POST requests. All state information is stored in the browser front end, allowing the back end to remain stateless, simplifying deployment. Including the Google Cloud FireStore API<sup>||</sup> enables the optional storage of results in a remotely hosted database. Plotting functionality is implemented using Chart.js<sup>\*\*</sup>

In common with the rest of SciKit-Surgery SciKit-SurgeryFRED utilises extensive testing and software process<sup>18</sup> to ensure the application is robust, reusable, and sustainable.<sup>26</sup> Change control and issue tracking is managed on GitHub<sup>††</sup> and continuous integration is managed using GitHub Actions.

§ <https://github.com/AIDanial/cloc> [v1.82]

¶ <https://palletsprojects.com/p/flask/> [v1.1.2]

|| <https://pypi.org/project/google-cloud-firestore/> [v2.0.1]

\*\* <https://www.chartjs.org/> [v2.9.4]

†† <https://github.com/UCL/scikit-surgeryfred>

## 2.3 Availability and Usage

For those who want to quickly try SciKit-SurgeryFRED, we currently maintain a running instance hosted at <https://scikit-surgeryfred.ew.r.appspot.com/>. This should be accessible from most modern web browsers.

If you want to run a locally hosted instance, or modify the code for your needs, you can download the source code. SciKit-SurgeryFRED is entirely open source software and is tested on Linux, MacOS, and Windows. The latest version can be obtained from Github. Alternatively, archived releases can be retrieved via Zenodo.<sup>20</sup>

```
1 git clone https://github.com/UCL/scikit-surgeryfred
```

Once installed the dependencies can be installed on a local virtual machine using tox. The application can then be run with as follows. This should output a web address that you can open in a browser to run the software.

```
1 cd scikit-surgeryfred
2 tox
3 source .tox/py37/bin/activate
4 python main.py
```

## 2.4 Extension to Anisotropic Errors

In many cases FLE cannot be properly modelled as an isotropic, independent, random variable. In the case of optical tracking systems for example<sup>30</sup> the errors normal to the camera plane are approximately 3 times those parallel to the camera plane. It is straightforward to implement such an anisotropic FLE in SciKit-SurgeryFRED and test the effect on registration outcomes. The following code snippet is taken from main.py. Line 66 defines the ratio of FLE in three directions. By default they are all equal.

```
57 @app.route('/getfle', methods=['POST'])
58 def getfle():
59     """
60     Returns values for fiducial localisation errors
61     Values are randomly selected from a uniform
62     distribution from 0.5 to 5.0 pixels
63     """
64     fle_sd = np.random.uniform(low=0.5, high=5.0)
65     #change fle_ratio if you want anisotropic fle
66     fle_ratio = np.array([1.0, 1.0, 1.0], dtype=np.float64)
67     anis_scale = math.sqrt(3.0 / (np.linalg.norm(fle_ratio) ** 2))
68     fixed_fle = fle_ratio * fle_sd * anis_scale
69
70     moving_fle = np.array([0., 0., 0.], dtype=np.float64)
71     fixed_fle_eavs = expected_absolute_value(fixed_fle)
72     moving_fle_eavs = expected_absolute_value(moving_fle)
73
74     returnjson = jsonify({
75         'fixed_fle_sd': fixed_fle.tolist(),
76         'moving_fle_sd': moving_fle.tolist(),
77         'fixed_fle_eav': fixed_fle_eavs.tolist(),
78         'moving_fle_eav': moving_fle_eavs.tolist()
79     })
80     return returnjson
```

Changing line 66 to;

```
66 file_ratio = np.array([3.0, 1.0, 1.0], dtype=np.float64)
```

sets the error in the x direction to 3 times that in the y and z directions.

## 2.5 Addition of Systematic Errors

A second significant source of error that is usually overlooked is the presence of systematic errors. For example when using a tracked pointer for fiducial localisation, any pointer calibration error will be added to all fiducial markers. Similarly some optical tracking systems can introduce a systematic error on the tracking markers.<sup>23</sup> SciKit-SurgeryFRED allows systematic error to be added to each fiducial. The fiducial localisation error is set within the JavaScript function `init_fles()` (defined in `static/main.js`) each time a new target is set.

```
496 /**
497  * Sets the global fiducial localisation error (FLE)
498  */
499 function init_fles() {
500     fetch("/getfle", {
501         method: "POST",
502     })
503     .then(resp => {
504         if (resp.ok)
505             resp.json().then(data => {
506
507                 let preOpFLEStdDev = data.moving_fle_sd;
508                 let intraOpFLEStdDev = data.fixed_fle_sd;
509                 let preOpFLEEAV = data.moving_fle_eav;
510                 let intraOpFLEEAV = data.fixed_fle_eav;
511
512                 let preOpSysError = [0.0, 0.0, 0.0];
513                 let intraOpSysError = [0.0, 0.0, 0.0];
514
515                 FLE = { preOpFLEStdDev, intraOpFLEStdDev,
516                       preOpFLEEAV, intraOpFLEEAV,
517                       preOpSysError, intraOpSysError };
518
519                 expectedFLEText.innerHTML = Math.round(Math.sqrt(FLE.intraOpFLEEAV)*100)
520                 /100;
521             });
522     })
523     .catch(err => {
524         console.log("An error occured setting fles", err.message);
525     });
526 }
```

By default there is no systematic error, (lines 512 and 513). We can add a systematic interoperative error at line 513 as;

```
513     let intraOpSysError = [1.0 * (Math.random() - 0.5),
514                           1.0 * (Math.random() - 0.5),
515                           1.0 * (Math.random() - 0.5)];
```

In this case the error is an isotropic uniform random variable, in the range -0.5 to 0.5. This error will be applied to all fiducial markers for a given registration.

## 3. RESULTS

### 3.1 SciKit-SurgeryFRED as a Teaching Aid

SciKit-SurgeryFRED (v0.0.3)<sup>21</sup> was used at our Medical Summer School in 2020 with a cohort of 5 students. Informal feedback indicated that it had improved their understanding of fiducial based registration. SciKit-SurgeryFRED also forms part of the basic augmented reality demonstration (BARD) which was a finalist at the MICCAI 2020 educational challenge<sup>‡‡</sup>.

<sup>‡‡</sup><https://miccai-sb.github.io/materials.html>





Figure 4: Results of 213 registrations using an anisotropic model of FLE

### 3.2 Extension to Anisotropic Errors

We simulated an anisotropic independent FLE, with FLE in the x direction being 3 times that in the y and z, as described in Section 2.4. Errors were scaled so that the expected absolute value of the FLE was the same as for the isotropic case. SciKit-SurgeryFRED was then used to perform at least 200 simulated registrations, the results of which are shown in Fig. 4.

It was apparent when performing the registrations that the majority of target registration error was in the x direction. However this is not communicated in the statistics of Figure 4. It would be an interesting extension exercise to use SciKit-SurgeryFRED to explore ways of communicating anisotropic errors during treatment.

### 3.3 Addition of Systematic Errors

We added a systematic FLE as an isotropic uniform random variable, in the range -0.5 to 0.5, as described in Section 2.5. This error will be applied to all fiducial markers for a given registration. We performed at least 200 simulated registrations using SciKit-SurgeryFRED, the results of which are shown in Fig. 5.

It is noticeable that average TREs are higher than in cases where there is no systematic error, while FRE remains similar. This is as expected as FRE will not account for systematic errors. This is a useful demonstration of this effect, though it might be more instructive to implement systematic errors in the game based method, see Section 2.1, to investigate the likely clinical impact of these systematic errors.

### 3.4 Statistical Significance

As discussed in Section 2 it is useful to perform tests of statistical significance on SciKit-SurgeryFRED's registration results. We used the function `stats.linregress` from SciPy<sup>27</sup> to perform a Wald Test against the null hypothesis that the slope is zero. For the results shown in Figures 2, 4, and 5 there was no significant relationship between actual TRE and either actual or expected FRE. There was a significant relationship between actual TRE and the expected TRE. All of these results are as expected.

Tests against the expected value of FLE and the number of fiducial markers are more problematic, as they are not continuous variables, which should be apparent when looking at the results charts. The reason for this





Figure 5: Results of 201 registrations with systematic error.

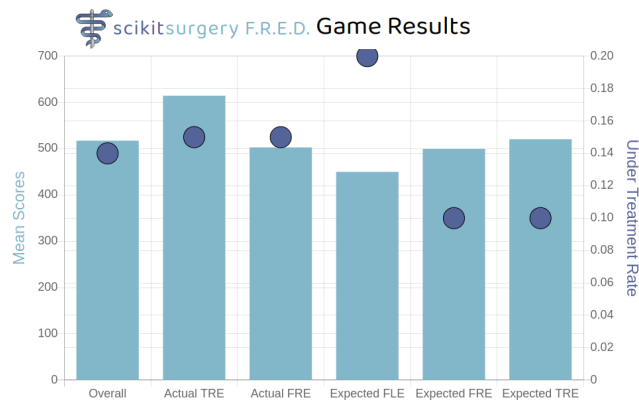


Figure 6: Average scores and treatment failure rates for the game based study.

is obvious for the number of fiducial markers. The expected value of FLE is clustered into groups as this is only set once for a given target, however the registration for this target will be repeated each time a new fiducial marker is added beyond the minimum of 3. Hence if the user adds a total of 10 fiducial markers, this will create 8 registration results all with different TRE and FRE, but with a single value of expected FLE.

### 3.5 SciKit-SurgeryFRED for Research

The results of the game based study described in Section 2.1 are shown in Figure 6. As expected scores are highest when the actual TRE is known. Interestingly it appears that when told only the expected value of the FLE the students tended to under treat the target more, resulting in lower overall scores.

The statistical significance of the results was tested using a two sided, unpaired, T-test implemented in SciPy's<sup>27</sup> `stats.ttest_ind` function. Using a p-value of 0.05 none of the results were statistically significant. Currently there are only 100 data points (20 for each category), so this is not surprising.

## 4. DISCUSSION

SciKit-SurgeryFRED has proven useful for the teaching of fiducial based registration, though more work on usability and measurement of teaching outcomes would be helpful. The results of the simulated ablation study are also of interest, however more participants are needed before any firm conclusions can be drawn. If we assume that additional data follow the same pattern as observed to date we would need to run the game on a further 15 subjects to show statistical significance on the difference in scores when using actual TRE. In the case of the FLE we would need a further 30 participants. Given the ease with which SciKit-SurgeryFRED can be deployed and used we hope to achieve this in the near future.

Currently SciKit-SurgeryFRED implements numerical measures for registration errors, however in practice it may be better to use graphical representations of registration error. An example is the use of outline rendering to show the misalignment of anatomical edges during keyhole surgery.<sup>24</sup> Azimi et al.<sup>3</sup> provide a similar example of a neuro-surgical guidance system where the registration error is communicated to the surgeon via misalignment of external anatomy. Such a method could be relatively easily implemented in SciKit-SurgeryFRED, with gamification used to measure its effectiveness.

Currently SciKit-SurgeryFRED is limited to rigid registration using point correspondence, however, there is no architectural reason why it could not be used to examine the effects of registration uncertainty more generally. It may be useful to implement more general error prediction models for rigid registrations.<sup>8,29</sup>

Surface based registration has long been proposed as a way to improve registration accuracy for neurosurgery<sup>15</sup> and is now integrated into commercial systems with mixed results.<sup>16</sup> SciKit-SurgeryFRED could be extended to include different registration methods, though obviously would require a more inclusive name. As we move to non rigid and probabilistic registrations, the correct interpretation of registration uncertainty will become more challenging.<sup>14</sup>

## 5. CONCLUSION

Understanding registration uncertainty is essential for the correct use of image guided surgery. There is, however, a lack of tools to aid teaching and research. We have presented SciKit-SurgeryFRED and demonstrated its use for teaching and research. SciKit-SurgeryFRED's ease of use enables it be deployed and demonstrated rapidly. SciKit-SurgeryFRED is open source and readily available so researchers and educators can modify and apply SciKit-SurgeryFRED based on their specific requirements.

## 6. OPEN ACCESS AND OPEN DATA

This research was funded in whole, or in part, by the Wellcome Trust [203145Z/16/Z]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. The registration data described in Sections 2 and 3 can be downloaded from <https://doi.org/10.5281/zenodo.4434278>.

## 7. ACKNOWLEDGEMENTS

This work is supported by the Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS) [203145Z/16/Z]. Our thanks to Simone Foti for designing the SciKit-Surgery logo, and to Dr Mary Thompson for proof reading the manuscript.

## REFERENCES

1. Anderson, L., Krathwohl, D., Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., Wittrock, M.: *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman (2001)
2. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-Squares Fitting of Two 3-D Point Sets. *IEE Transactions on Pattern Analysis and Machine Intelligence* (5), 698–700 (1987)

3. Azimi, E., Niu, Z., Stiber, M., Greene, N., Liu, R., Molina, C., Huang, J., Huang, C.M., Kazanzides, P.: An interactive mixed reality platform for bedside surgical procedures. In: A.L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M.A. Zuluaga, S.K. Zhou, D. Racoceanu, L. Joskowicz (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 65–75. Springer International Publishing, Cham (2020)
4. Chan, W.Y., Qin, J., Chui, Y.P., Heng, P.A.: A serious game for learning ultrasound-guided needle placement skills. *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society* **16** (2012). URL <https://doi.org/10.1109/TITB.2012.2204406>
5. Clarkson, M., Dowrick, T., Thompson, S., Ahmad, M., Koo, B.: *scikit-surgerycore: v0.6.6 summer school 2020* (2020). URL <https://doi.org/10.5281/zenodo.3965731>
6. Collins, K., Kanev, K., Kapralos, B.: Using games as a method of evaluation of usability and user experience in human-computer interaction design. In: *Proceedings of the 13th International Conference on Humans and Computers*, pp. 5–10 (2010)
7. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses: The species severe acute respiratory syndrome-related coronavirus: classifying 2019-ncov and naming it sars-cov-2. *Nature microbiology* **5**(4), 536–544 (2020). URL <https://doi.org/10.1038/s41564-020-0695-z>
8. Danilchenko, A., Fitzpatrick, J.M.: General approach to first-order error prediction in rigid point registration. *IEEE Transactions on Medical Imaging* **30**(3), 679–693 (2011). URL <https://doi.org/10.1109/TMI.2010.2091513>
9. Fitzpatrick, J., West, J., Maurer C.R., J.: Predicting error in rigid-body point-based registration. *Medical Imaging, IEEE Transactions on* **17**(5), 694–702 (1998). URL <https://doi.org/10.1109/42.736021>
10. Fitzpatrick, J.M.: Fiducial registration error and target registration error are uncorrelated. In: M.I. Miga, K.H. Wong (eds.) *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, vol. 7261, pp. 21 – 32. International Society for Optics and Photonics, SPIE (2009). URL <https://doi.org/10.1117/12.813601>
11. Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E.: Array programming with NumPy. *Nature* **585**, 357–362 (2020). URL <https://doi.org/10.1038/s41586-020-2649-2>
12. Hunt, A., Thomas, D.: *The Pragmatic Programmer: From Journeyman to Master*. Addison-Wesley Longman Publishing Co., Inc., USA (2000)
13. Katz, D., Zerillo, J., Kim, S., Hill, B., Wang, R., Goldberg, A., DeMaria, S.: Serious gaming for orthotopic liver transplant anesthesiology: A randomized control trial. *Liver transplantation : official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society* **23**(4), 430–439 (2017). URL <https://doi.org/10.1002/lt.24732>
14. Luo, J., Frisken, S., Wang, D., Golby, A., Sugiyama, M., Wells III, W.: Are registration uncertainty and error monotonically associated? In: A.L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M.A. Zuluaga, S.K. Zhou, D. Racoceanu, L. Joskowicz (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 264–274. Springer International Publishing, Cham (2020)
15. Maurer, C.R., Maciunas, R.J., Fitzpatrick, J.M.: Registration of head CT images to physical space using a weighted combination of points and surfaces. *IEEE Transactions on Medical Imaging* **17**(5), 753–761 (1998)
16. Mongen, M., Willems, P.: Current accuracy of surface matching compared to adhesive markers in patient-to-image registration. *Acta Neurochirurgica* **161** (2019). URL <https://doi.org/10.1007/s00701-019-03867-8>
17. Schneider, C., Thompson, S., Totz, J., Song, Y., Allam, M., Sodergren, M., Desjardins, A., Barratt, D., Ourselin, S., Gurusamy, K., Stoyanov, D., Clarkson, M., Hawkes, D., Davidson, B.: Comparison of manual and semi-automatic registration in augmented reality image-guided liver surgery: a clinical feasibility study. *Surgical endoscopy* **34**(10), 4702–4711 (2020). URL <https://doi.org/10.1007/s00464-020-07807-x>

18. Schroeder, W.J., Ibanez, L., Martin, K.M.: Software process: the key to developing robust, reusable and maintainable open-source software. In: 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), pp. 648–651 Vol. 1 (2004). URL <https://doi.org/10.1109/ISBI.2004.1398621>
19. Siewerdsen, J., Adrales, G., Anderson, W., Carey, J., Creighton, F., Dibrito, S., Galaiya, D., Marohn, M., McNutt, T., Osgood, G., Theodore, N., Weiss, C., Viswanathan, A.: Surgineering: curriculum concept for experiential learning in upper-level biomedical engineering. *International Journal of Computer Assisted Radiology and Surgery* **15** (2019). URL <https://doi.org/10.1007/s11548-019-02094-x>
20. Thompson, S.: Ucl/scikit-surgeryfred: v0.1.3 spie medical imaging 2021 (2021). URL <https://doi.org/10.5281/zenodo.4462897>
21. Thompson, S., Ahmad, M.: scikit-surgeryfred: v0.0.3 summer school 2020 (2020). URL <https://doi.org/10.5281/zenodo.3946090>
22. Thompson, S., Dowrick, T., Ahmad, M., Xiao, G., Koo, B., Bonmati, E., Kahl, K., Clarkson, M.J.: Scikit-surgery: compact libraries for surgical navigation. *International journal of computer assisted radiology and surgery* **15**(7), 1075–1084 (2020). URL <https://doi.org/10.1007/s11548-020-02180-5>
23. Thompson, S., Penney, G., Dasgupta, P., Hawkes, D.: Improved modelling of tool tracking errors by modelling dependent marker errors. *IEEE Transactions on Medical Imaging* **32**(2), 165–177 (2013). URL <https://doi.org/10.1109/TMI.2012.2216890>
24. Thompson, S., Schneider, C., Bosi, M., Gurusamy, K., Ourselin, S., Davidson, B., Hawkes, D., Clarkson, M.J.: In vivo estimation of target registration errors during augmented reality laparoscopic surgery. *International journal of computer assisted radiology and surgery* **13**(6), 865–874 (2018). URL <https://doi.org/10.1007/s11548-018-1761-3>
25. Tofade, T., Elsner, J., Haines, S.T.: Best practice strategies for effective use of questions as a teaching tool. *American Journal of Pharmaceutical Education* **77**(7) (2013). URL <https://doi.org/10.5688/ajpe777155>
26. Venters, C.C., Capilla, R., Betz, S., Penzenstadler, B., Crick, T., Crouch, S., Nakagawa, E.Y., Becker, C., Carrillo, C.: Software sustainability: Research and practice from a software architecture viewpoint. *Journal of Systems and Software* **138**, 174 – 188 (2018). URL <https://doi.org/10.1016/j.jss.2017.12.026>
27. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17**, 261–272 (2020). URL <https://doi.org/10.1038/s41592-019-0686-2>
28. West, J.B., Maurer, C.R.: Designing optically tracked instruments for image-guided surgery. *IEEE Transactions on Medical Imaging* **23**(5), 533–545 (2004). URL <https://doi.org/10.1109/TMI.2004.825614>
29. Wiles, A.D., Likholyot, A., Frantz, D.D., Peters, T.M.: A statistical model for point-based target registration error with anisotropic fiducial localizer error. *IEEE Transactions on Medical Imaging* **27**(3), 378–390 (2008). URL <https://doi.org/10.1109/42.736031>
30. Wiles, A.D., Thompson, D.G., Frantz, D.D.: Accuracy assessment and interpretation for optical tracking systems. In: R.L.G. Jr. (ed.) *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*, vol. 5367, pp. 421 – 432. International Society for Optics and Photonics, SPIE (2004). URL <https://doi.org/10.1117/12.536128>