Journal of Health and Quality of Life

Journal homepage:
https://karyailham.com.my/index.php/jhqol/index
ISSN: 3030-5101


# Validation of Glycemic Index Values Represented as Bar Graphs in Scholarly Articles for Clinical Nutrition Application

Bhuwaneswaran Vijayam[1,2,*], Lau Siew Ping[1], Wong Jing Wen[1], Sakinah Muzarpar[1], Ainun Amirah Abdul Rahim[1], Pwi Jia Yee[1], Leeynesh Sooriyapiragasam[1]

[1] Newcastle University Medicine Malaysia, 79200 Iskandar Puteri, Johor, Malaysia
[2] Johor Bahru Parenteral and Enteral Nutrition Society (JPENS), 81000 Johor Bahru, Johor, Malaysia

**ARTICLE INFO**

***Article history:***
Received 28 August 2023
Received in revised form 25 November 2023
Accepted 28 November 2023
Available online 26 January 2024

***Keywords:***

Glycemic index; Glycaemic index; ImageJ bar; Bar graph analysis; Precision nutrition; Clinical nutrition

**ABSTRACT**

Glycemic index (GI) has been a proven clinical nutrition tool. It has been utilized in chronic diseases such as PCOS, diabetes, obesity and metabolic syndrome. The GI value of a test food is the ratio of the glucose response over 2 hours against the reference food. The references food is usually either glucose or bread. GI values may often be directly available as values in texts, tables, or figures. The indirect values may be represented as values, images or both values and, concentration of glucose over time or GI values. Less ideally, some GI data are only presented in graphs. Data from these graphs without direct values could be extracted but this potentially contributes to less accurate results. Here, we investigate the extracted outcome from the original data and if their differences could affect the GI values. There are a few software's and methods that may be used for data extraction; however, they are costly and complex. Also, different types of graphs require different extraction methods. Herein, we describe a simple reproducible method for extracting data from bar graphs using the freeware ImageJ. Seven extractors extracted 102 outcome values from 19 different scholarly articles. Differences between extractors were compared using the Overall Concordance Correlation Coefficients (OCCC), whereas differences between the original and extracted data were compared using the Lin's Concordance Correlation Coefficients (CCC) and Bland-Altman. Both the OCCCs and CCCs were high for both outcome values and errors, while the CCCs were good and acceptable for the outcome. The Bland Altman showed good agreement between the extracted outcome and the true reported GI values. Therefore, the current method supports the extraction GI values from bar graphs in published scholarly articles even in their absence in text.

## 1. Introduction

It is predicted that the prevalence of Diabetes Mellitus (DM) will increase by more than 50% by 2030. We are now less than a decade away from fulfilling this scientific forecast by Rowley *et al.,* [1]. As the treatment options of diabetes are getting more advanced and personalized by the day,

* *Corresponding author.*
*E-mail address: Bhuwaneswaran.Vijayam@newcastle.edu.my*

https://doi.org/10.37934/hqol.1.1.115

utilization of glycemic index (GI) has been thoroughly discussed as an additional measure by Zafar *et al.,* and Unwin *et al.,* in two different manuscripts [2,3]. In clinical practice, GI has been proven to have an impact in the rehabilitation and treatment of chronic conditions such as central obesity, DM, prediabetes, Polycystic Ovarian Syndrome (PCOS) and metabolic syndrome [2,4-6].

According to World Health Organisation (WHO), GI is the value derived from the calculation of periodic blood glucose response after consumption of a specific food. To be precise, Food and Agriculture Organisation (FAO) has defined GI as the incremental area under the curve (iAUC) for the aforementioned blood glucose response of the test food when compared or indexed to a reference food [8]. This iAUC neglects the area under the fasting point as proposed by Brouns *et al.,* [9]. Typically, white bread and glucose have been used exclusively as the reference food in most clinical nutrition articles for GI calculation as reported by three different authors in their respective compilation [10-12].

Along the years, our team have encountered numerous GI manuscripts that have represented these values as direct values, bar graphs, box plots, glucose response values, iAUCs and glucose response curves [13-15]. Although plot figures are undeniably attractive, clear, and concise way to demonstrate GI values, some scholarly articles solely report them without the exact values in text. These articles are less likely to be discerned by other researchers and clinicians alike as the GI values based upon visualizing the plot figures alone could not be inferred. This has limited the potential of the scholarly articles to expand the large data pool of GI for the translation of research into clinical nutrition application in the real world. As this was an apparent clinical gap that our team has encountered while deciphering GI bar graphs, we decided to test out and propose a solution to this issue so that these valuable data can be utilized optimally.

In general, according to Bajic *et al.,* data represented as figures could be either automatically or interactively extracted [16]. These could be carried out via photo editing programs, image processing programs or custom coded algorithms. In this research, to the capacity of our clinical experience, we would be utilizing ImageJ. ImageJ is a Java based image analysis program which has shown paramount value in many scientific projects particularly in biological sciences [17]. ImageJ is developed by the research service of the National Institute of Health (NIH; Bethesda, MD, USA) and released in 1997 [17,18]. In addition, ImageJ is free, expandable, and available for all operating systems [17].

The aim of this research is to extract GI values represented as bar graphs in selected published scholarly articles using ImageJ. The extracted values would then be compared to the reported value in the corresponding scholarly articles. The values are then further analysed to evaluate accuracy, precision, and agreement.

## 2. Methodology

This research was conducted from August 2023 by a team of individuals consisting of seven academic staffs in Newcastle University Medicine Malaysia (NUMed). The seven extractors were BV, LS, LSP, WJW, SM, AAAR and PJY. A brief introduction on the usage of ImageJ was given by BV beforehand as the rest of the members did not have previous experience with the said application. Google images with GI were pooled by typing the keywords "glycemic index" and "glycaemic index bar". Published scholarly articles were retrieved according to the respective Google images. All images from the scholarly articles were pooled only after meeting their inclusion criteria and exclusion criteria. The inclusion criteria were all articles that had both GI represented as bar graphs and GI values embedded in the scholarly article's text, table, line curves or at any part of the bar graph. Scholarly articles with appropriate and suitable bar graphs representing the GI were pooled for this research. Cases where difficulties would be anticipated during GI extraction due to image, bar, axis, or scale factors were re-examined and excluded after consensus amongst the extractors. All the pooled scholarly articles were saved as portable document format (.PDF) files. These aside, bar graphs that represented the iAUC values of glucose instead of GI were excluded as they required additional calculation steps to obtain their GI values. The overall process is summarized in Figure 1.

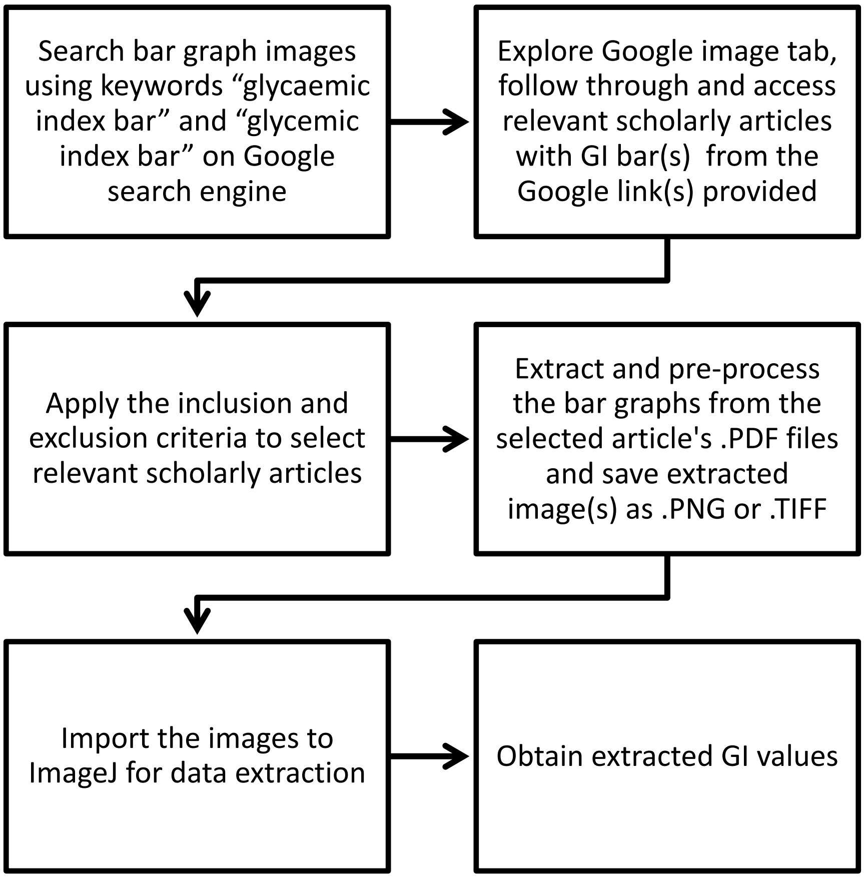


**Fig. 1.** Overall methodology for this research

### *2.1 Image Pre-Processing*

The bar graph from each scholarly article was extracted and saved as Portable Network Graphic (.PNG) or Tag Image File Format (.TIFF) format to preserve pixel integrity and reduce aliasing. Some of the bar graphs were pre-processed to remove values that were displayed in the respective images or bars [19-26].

### *2.2 GI Extraction Via ImageJ*

For this research, all extractors utilized monitor screens with the resolution of at least 1920 × 1080 pixels. ImageJ was obtained from https://imagej.nih.gov/ij/download.html. The calculation algorithm utilized by ImageJ's 'length' and "Measure" function is the calculation of pixel distance and scaling it to the known value or "real-life" value of the scale. Hence the scaling method for GI values extracted via ImageJ ($GI_{\text{extracted}}$) calculation was done using the equation given. $GI_{\text{Bar length}}\,TF$ is the distance/ or length in pixels of test food bar while $GI_{\text{Bar length}}\,RF$ is the pixel of reference food bar or y-scale axis representing the GI scale.

$$\text{GI } extracted = \frac{\text{GI Bar } length\ TF}{\text{GI Bar } length\ RF} \times 100$$

True reported values of GI were extracted from either the text, table or the graph of each scholarly article. For illustration purpose of this article, we created a mock GI bar graph which is similar to those acquired from scholarly article PDFs as shown in Figure 2(a). Ideally, the image acquired from the PDF could be pre-processed to optimize the size and workspace prior to importing into ImageJ. Firstly, the .PNG or .TIFF image file is loaded in ImageJ as shown in Figure 2(b). This was done via the “File> Open”. The scale was set using the glucose bar as reference. This was done after placing the ‘line selection tool’ according to the height of the reference food which is the glucose bar. The scale is set after accessing the dialog box to change the known distance via “Analyse > Set Scale”. These are depicted in Figure 2(c) and Figure 2(d). Similarly, this could be done on the y-axis, and to measure the GI scale in the absence of reference food. Next, this process was repeated for a test food, in this case milk, as shown in Figure 2(e). Upon placement of another ‘line selection tool’, “Measure” was accessed via the “Analyse” tab. The extracted GI value was displayed as ‘length’ (adjacent to ‘angle’) as depicted in Figure 2(f). The extracted GI value is derived via the mathematical equation as previously described in Eq. (1). All these steps with ImageJ are depicted in Figure 2.

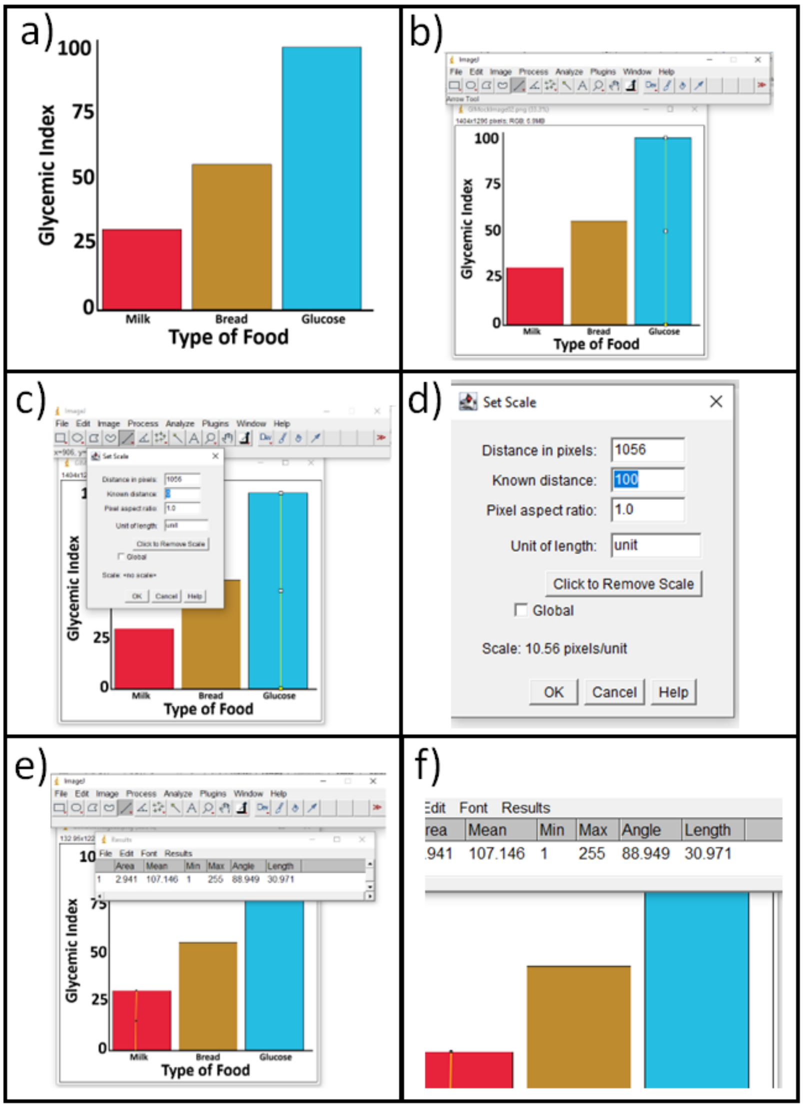


**Fig. 2.** Steps for GI extraction via imageJ

*2.3 Outcome Measures*

GI extracted from ImageJ were compared with the true reported values to obtain their differences, if any. The use of integrals or decimal places for each value were the same as their respective scholarly article to ensure consistency of the data. The time taken to measure each bar was recorded separately using a stopwatch. The computation time was recorded upon loading an image till the completion of GI data extraction process for each food. The computation time was self-reported by each extractor. The selected scholarly articles were randomly distributed amongst the extractors for analysis. Data tabulated by each extractor was crosschecked by another member from the team.

*2.4 Data Analysis*

Data were tabulated and analysed in Microsoft Excel and MedCalc® Statistical Software version 22.009 (MedCalc Software Ltd, Ostend, Belgium; https://www.medcalc.org; 2023). Data normality and goodness were tested with Shapiro Wilk and D'Agostino Pearson test for small and large dataset respectively. The latter was chosen for large dataset as the said test has good performance in term of power [27]. The values were represented as either mean and standard deviation (SD) or median and interquartile range (IQR) based on the normality of the dataset. The former combination was chosen for normally distributed data.

A few Likert scales were developed in accordance with Sullivan *et al.,* [28]. A scale with scores 1-5 (1: Never, 2: Rarely, 3: Sometimes, 4: Often and 5: Always) was used to evaluate each extractor's experience with ImageJ. Another similar scale was used to evaluate average computer usage (1: Never, 2: Rarely, 3: Sometimes, 4: Weekly and 5: Almost Daily). The last scale also consisted of 5 choices (1: Very Difficult, 2: Difficult, 3: Fair, 4: Easy and 5: Very Easy) to evaluate each extractor's ease of utilizing ImageJ.

Lin's Concordance of Correlation Coefficient (CCC) was used to evaluate the concordance, precision ($P$) and accuracy ($X^a$) of the GI values extracted from ImageJ when compared to the true reported GI values [29]. The Lin's CCC measurement indicates how well a new set of observation (GI extraction from bar graphs) reproduces the true GI values. Similarly, the individual concordance, $P$ and $X^a$ data for each extractor were analysed and the mean was represented as Overall CCC (OCCC) as proposed by Banhart [30].

A Bland-Altman (BA) was used to test biasness, limit of agreement and coefficient of repeatability [31].

To test differences of efficiency analysis, an analysis of co-variance (ANCOVA) was utilized to test the differences of computation time amongst extractors. This test was also utilized for other nominal measurements when suitable such as GI value differences. A Kruskal-Wallis test was utilized to report the differences between ordinal values and the extractors' values. The extractors were assigned as the covariate and code factor for the ANCOVA and Kruskal-Wallis tests respectively.

All values were reported to the closest three decimal places. A *p-value* of <0.05 was considered statistically significant.

## 3. Results

The number of data analysed in this research was 102, pooled from 19 selected scholarly articles. The scholarly articles' publications dates ranged from the year 2004 to 2022. The types of bar graphs encountered by the extractors were all basic vertical plots. Some food GIs were represented solely as bar graph without values in text. These bar graphs were not included for the extraction process. The characteristic of the bar graph images, and the GI values pooled are summarized in Table 1. As previously stated, some images were pre-processed to remove GI values from the bar, prior to analysis in ImageJ to reduce interference and biasness to the extraction process by extractors.

**Table 1**
The characteristics and properties of the bar graphs and GI of the scholarly articles analysed

| Author | Year | Type of Food | Number of Images Extracted | Bar graph Properties | True GI Value Placement | True GI Value Decimal Place(s) | Reference |
|---|---|---|---|---|---|---|---|
| Akinjayeju *et. al.,* | 2020 | Protein Maize, Soy Cake, Whole Pearl Millet Flours | 3 | Basic Vertical | In text | Integer | [32] |
| Akinyede *et. al.,* | 2022 | Raw Malabar Chestnut Seeds, Cooked Malabar Chest Seeds, Roasted Malabar Chestnut Seeds | 3 | Basic Vertical | Bar graph | Two | [22] |
| Almousa *et. al.,* | 2013 | Rigag, Shbab, Khameer, White Pita Bread (Khobuz), White Bread Roll (Summon), Wholemeal Bread Slices | 6 | Basic Vertical | Bar graph | Two | [24] |
| Assefa *et. al.,* | 2017 | White Teff Enjara, Red Teff Enjera, Maize Enjera, Barley Bread, Maize Bread, Wheat Bread, Qoch'o Bread, White Bread, Bullo genfo, Pea sauce, Chickenpea sauce, Lentil sauce | 12 | Basic Vertical | In text | Integer | [33] |
| Blair *et. al.,* | 2006 | Chocolate Daycream Sucralose Shake, Chocolate Daycream Fructose Shake, Soy Spaghetti, Chocolate Raspberry Zing Bar, Peanut Butter Chocolate Bar, Soy Protein Chips (Lightly Salted) | 6 | Basic Vertical | Table | Two | [34] |
| Bornet *et. al.,* | 1987 | Bread, Potato, Spaghetti, Rice, Lentils, Beans | 6 | Basic Vertical | In text | Integer | [35] |
| Caballero-de la Peña *et. al.,* | 2022 | Mango drink with Soy/Maize Protein with Sucrose, Mango drink with Soy/Maize Protein with Stevia & Sucralose, Mango drink with Whey Protein Concentrate, Mango drink, Stevia | 5 | Basic Vertical | Bar graph | Integer | [25] |
| Chatuverdi *et. al.,* | 2017 | Chakli (5% Kale), Twister (10% Kale) | 2 | Basic Vertical | Bar graph | Two | [23] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Di Cairano *et. al.,* | 2022 | Control biscuit from Sucrose & without Resistant Starch, Biscuit with 12% Resistant Starch & 30% Insulin replacing Sucrose, Biscuit with 12% Resistant Starch & 50% Maltitol replacing Sucrose, Biscuit with Total Replacement of Sucrose by 100% Maltitol | 4 | Basic Vertical | Table | Integer | [36] |
| Eldakhakhny *et. al.,* | 2021 | Sucrose (Male), Sucrose (Female), Sucrose (Combined) | 3 | Basic Vertical | Line graph | Integer | [21] |
| Flint *et. al.,* | 2004 | Finnish bread with Butter & Cheese,<br>German bread with Butter and Cheese.<br>Reference bread Butter & Cheese,<br>Italian biscuits with Coffee & Milk,<br>Reference bread with Butter,<br>All-bran plus and Milk,<br>Reference bread with Butter & Jam,<br>Rolled oats with Sugar & Milk,<br>Frosties Milk,<br>All-bran with Milk,<br>French bread with Butter and Jam,<br>Cornflakes & Milk,<br>Rolled oats Porridge with water & Apple sauce | 13 | Basic Vertical | Bar graph | Integer | [20] |
| Geetha *et. al.,* | 2020 | Roti,<br>Dosa,<br>Dumpling | 3 | Basic Vertical | Bar graph | Integer | [26] |
| Jimoh *et. al.,* | 2008 | Boiled Yam,<br>Pounded Yam,<br>Amala Yam | 3 | Basic Vertical | Table | One | [37] |
| Luke *et. al.,* | 2018 | Sprite™, Sprite™ + 2.6g BTI320, Sprite™ + 5.2g BTI320 | 3 | Basic Vertical | In text | Integer | [38] |
| Nagaraju *et. al.,* | 2020 | C1 Multigrain Indian Bread, C2 Multigrain Indian Bread | 2 | Basic Vertical | In text | One | [39] |
| Olugbuyi *et. al.,* | 2021 | 100% Ceroline Dough Meal (CERD), Optimized Dough Meal [Plantain 60%, Soycake 25%, Rice Bran 15%] (PSRD) | 2 | Basic Vertical | In text | Integer | [40] |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Oluwajuyitan *et. al.,* | 2019 | Formulated Dough meals - 100% Plantain (PLT), 59.83% Tigernut with 40.17% Defatted Soybean (TNS), 100% Commercial Dough meal (CNT) | 3 | Basic Vertical | In text | Two | [41] |
| Origbemisoye *et. al.,* | 2021 | Cookies with 20g Margarine + 10g Ackee Arils Flour (AF1), Cookies prepared from 30g Ackee Arils Flour (AF3), Cookies with 30g Margarine (NAFC) | 3 | Basic Vertical | In text | Two | [42] |
| Rackzkowska *et. al.,* | 2019 | Traditional Steamed Dumplings, Traditional Pancakes with Apple, Modified Rolls with Cheese, ModifieD Pancakes with Apple | 4 | Basic Vertical | In text | One | [43] |
| Zhao *et. al.,* | 2023 | Multiple formulations of low GI and gluten-free biscuits using Potato Flour, Indica Rice Flour, Oat Bran and Inulin | 16 | Basic Vertical | Bar graph | Two | [19] |

The median age of the seven extractors was 25 (±3.225) years. The highest education level of all extractors was at least a bachelor's degree. All the extractors used computer daily and almost all (except for one extractor) of them had no experience utilizing ImageJ. In general, there were no statistical differences among extractors for average computer usage, experience with ImageJ, and computation time taken to extract GI. There were also no differences between extracted GI values and the true reported values amongst extractors. Table 2 shows the numerical representation of the previously narrated data.

**Table 2**
Characteristics differences between extractors

| Parameter | Mean (SD) or Median (IQR)* or Likert description** | Normality Test | ANCOVA F(*df1,df2*)=F Ratio$^{\$}$ or Kruskal-Walis Chi Square,*df*$^{\#}$ | *p-value* |
|---|---|---|---|---|
| Extractors average computer usage | 5 (0)** | reject normality (p<0.000) | 2.250,6$^{\#}$ | *0.423* |
| Extractors experience with ImageJ | 1 (0)** | reject normality (p<0.000) | 2.25,6$^{\#}$ | *0.423* |
| Computation time per GI bar | 10.560 (4.8)* | reject normality (p<0.000) | F(1,100) = 0.025$^{\$}$ | *0.825* |
| Difference of extracted GI and true reported GI values | 0.175 (0.399)* | reject normality (p<0.000) | F(1,100) = 0.174$^{\$}$ | *0.677* |

*Journal of Health and Quality of Life*
Volume 1, Issue 1 (2024) 1-15

The individual average computation time and differences of extracted GI values from the true reported values according to extractors are depicted side by side in different colours and scales in Figure 3. Values that were plotted as median are represented with an orange star symbol.

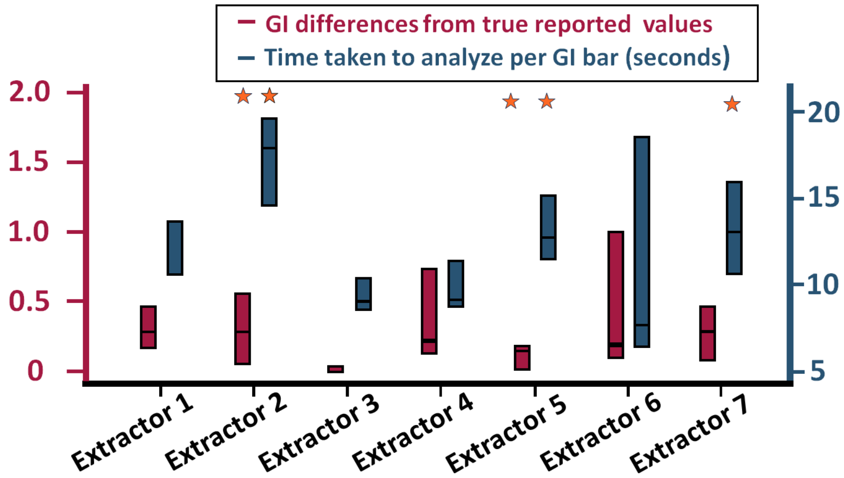


**Fig. 3.** The average extracted GI differences and time taken to extract GI data among extractors

As for the CCCs, the Lin's CCC was 0.9994 (CI: ±0.0002). The OCCC was 0.999 as projected in Table 3.

**Table 3**
The Lin's CCC of individual extractors and the OCCC

| Extractor | CCC | *P* | $X^a$ |
|---|---|---|---|
| Extractor 1 | 0.999 | 0.999 | 1 |
| Extractor 2 | 0.999 | 0.999 | 1 |
| Extractor 3 | 0.999 | 0.999 | 1 |
| Extractor 4 | 0.998 | 0.998 | 1 |
| Extractor 5 | 1 | 1 | 1 |
| Extractor 6 | 0.999 | 0.999 | 1 |
| Extractor 7 | 0.999 | 1 | 1 |
| **Overall** | **0.999** | **0.999** | **1** |

In addition, the BA analysis synthesized a biasness of 0.006, an upper Limit of Agreement (LoA) of 1.119, and a lower LoA of -1.107. The previously described BA plot is represented in Figure 4. The blue line and two dotted red lines represent the biasness and LoAs respectively. Furthermore, 97.1% data points lie within the LoA. The BA coefficient of repeatability was 1.108.

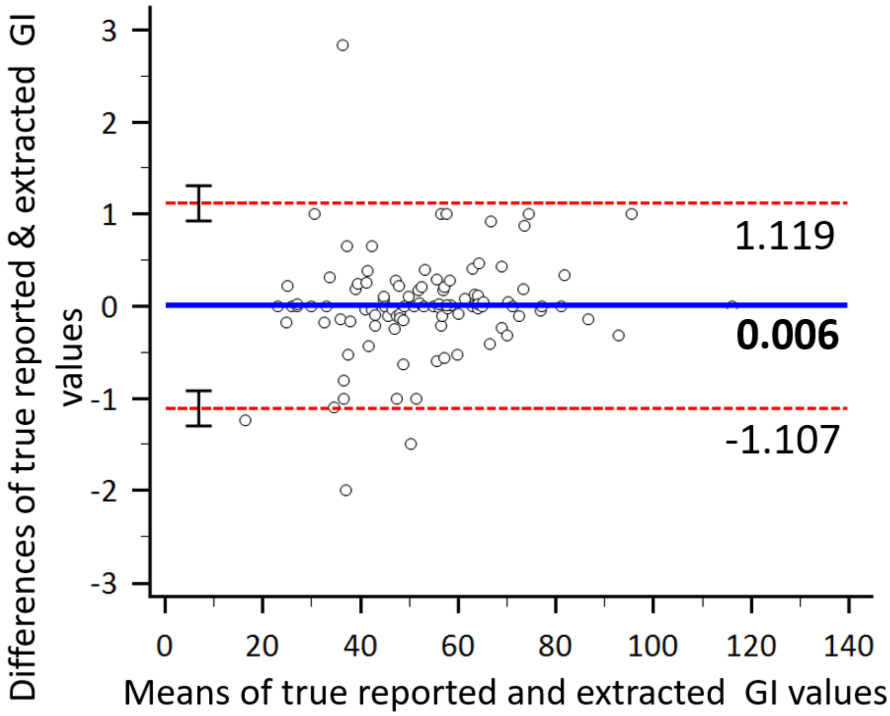


**Fig. 4.** Bland-Altman plot of extracted GI values and true reported GI values

Lastly, the ease of ImageJ usage according to the extractors was found to be 4 (Easy). There were no significant differences to this response amongst extractors Kruskal-Walis (Chi Square, *df* )= (3.75 , 6 ), p=0.423.

## 4. Discussion

In this research, we mainly investigated the concordance between GI values extracted from bar graphs and the true reported GI values in scholarly articles. The concordance of GI value extracted by each extractor and the differences between extractors were also investigated. McBride proposed that CCC values of 0.95-0.99 as substantial, and >0.99 as almost perfect strength of agreement [44]. Our results demonstrated an almost perfect strength-of-agreement criteria for both Lin's CCC and OCCC. This suggests that the current GI extraction method via ImageJ has a statistically defensible measure of concordance and significantly high accuracy. In addition, our analysis shows that the data extracted from this method have great clinical reliability due to its low BA coefficient of repeatability as previously discussed by Altman in two separate research [31,45]. Therefore 95% of differences between repeated GI extraction with ImageJ are expected to be within 1.1. Elsewhere Gheibi *et al.,* reported similar CCC and BA findings when extracting data from bar graphs with Adobe Photoshop [46].

Notably, the CCC, accuracy and precision of the GI values extracted were high given that 39% of our lines that were placed by all extractors were not perfect straight lines and had some degree of angles (data not shown). The accuracy and precision could further be improved with cautious, diligent plotting and the subsequent reduction of the previously mentioned angles. Although it is generally presumed that this could also be improved with automated data extraction, real life data shows otherwise. Recent research by Rane *et al.,* utilizing automated data extraction from bar graphs had an accuracy ranging between 17-77% [47]. Another research by Al-Zaidy *et al.,* which also utilized automated data extraction had a higher accuracy ranging between 77-95% [48].

Otherwise, we solely encountered GI values that were represented as vertical bars from the 19 scholarly articles analysed in this research. Typically, bar graphs could be divided into basic, grouped, or stacked bars represented either vertically or horizontally as suggested by Bajic *et al.,* and Mishra *et al.,* in their respective research [16,49]. Based on this premise, we deduce that ImageJ could also be potentially utilized for data mining in linear plot graphs, box plots, 3-Dimensional (3D) bar graphs that are represented as single, grouped, or stacked graphs.

To the best of our knowledge, this is the first paper to extract, validate and discuss GI data from bar graphs published in scholarly articles. This research is, inter alia, essential as future research could utilize this method to extract GI values from scholarly articles and expand the clinical nutrition database. A recent database reported over 4000 food GI values [50]. In the southeast Asian region, Henry *et al.,* and Osman *et al.,* have 940 and 83 food GI values in their research respectively [10,11]. All three of these databases have reported most of the GI values as an integer and occasionally at two decimal places. From a clinical nutrition point of view, the utilization of our GI extraction method could potentially skew the GI values by merely -1.1 to 1.1. At this point, this method can be considered precise and practical to be used in future research. This would aid data acquisition and synthesis in clinical nutrition systemic reviews and meta-analyses. With abundance of GI data, a low GI strategy could be mooted to combat nutrition related diseases. There are growing interests in utilizing nutritional means such as GI to lower hyperglycaemia, adiposity, blood pressure, insulin resistance, endothelial dysfunction, inflammation, β-Cell dysfunction and increased prothrombic factors [3,51,52].

The time taken to extract data from graphs vary according to methods and programs. Previous studies of this efficiency had a wide range of between 0.209 seconds and 900 seconds [47,53,54]. Given that the method we utilized involved human extractors with the aid of a program, the computation time in our research was indifferent from Moeyat *et al.,* [54]. Moeyart *et al.,* reported a range of 15-17 seconds with three different programs [54]. The shortest time taken was reported by Luo *et al.,* in which an automated deep hybrid framework of deep learning was used instead of humans [53]. The real-life translation of this method proposed by us would consume an average 11 seconds to extract a GI data from bar graphs. Furthermore, all our seven extractors had no statistical differences from one another in extracting GI values.

For this research, ImageJ program was chosen for the GI data extraction as it is a free, robust tool for both research and clinical nutrition. Our findings also showed that ImageJ could be comfortably handled by the extractors even for those without prior experience, as reported by the extractors themselves. At the point of writing this manuscript our team were only aware of a few other similar freeware namely Engauge Digitizer and WebPlotDigitizer applied by two separate research [55,56]. However, we have not tested these programs.

There are few limitations to this research. Firstly, this current GI extraction method was only tested with two-dimensional bar graphs. We did not test this method on complex three-dimensional (3D) bar images. Secondly, we did not extract any iAUC values represented as bar graphs to calculate and validate the GI values. As for ImageJ, the program does not allow 'assigning' and renaming the 'length' data to a variable of interest. The length data is manually tabulated in Microsoft Excel. As there are display of mean and other data on ImageJ, a tendency for human error is present during tabulation. This is time consuming for other secondary analyses in programs such as Microsoft Excel and Medcalc. Moreover, user memory recall is needed in multiple bars instances and the variable name is non-modifiable. This may be an issue for images with multiple bars and similar food names.

On top of that, this method was not tested for bar graphs captured from real-world photography, mobile phones, and computer screens. These images are sometimes misaligned such as tilted, skewed, de-pixelated, aliased, out of perspective, may have artefacts and distortions. As they may need more than image preprocessing and correction(s), programs such as ImageJ might not be suitable for direct data extraction as this could yield results with compromised accuracy.

Further work on this topic and the extraction method should involve a larger dataset with more diverse content, including line graphs, calculations of iAUCs, 3D-bar graphs, and box plots. On a broader perspective for clinical nutrition, our findings could be utilised to extract other nutrition indices such as the uric acid index and insulinaemic index (II) values that are represented as bar plots. Of note, this could also potentially be used in other biomedical fields when bar images are used to represent certain values.

## 5. Conclusion

In conclusion, the extraction of GI values in bar graphs via ImageJ is simple and reliable. The method also supports the extraction GI values from bar graphs in published scholarly articles even in their absence in text. This could potentially be expanded to other similar nutritional indices represented as bar graphs in the clinical nutrition field.

**Acknowledgement**
This research was funded by Newcastle University Medicine Malaysia NIRG 226900-IG9.

## References

[1] Rowley, William R., Clement Bezold, Yasemin Arikan, Erin Byrne, and Shannon Krohe. "Diabetes 2030: insights from yesterday, today, and future trends." *Population health management* 20, no. 1 (2017): 6-12. https://doi.org/10.1089/pop.2015.0181

[2] Zafar, Mohammad Ishraq, Kerry E. Mills, Juan Zheng, Anita Regmi, Sheng Qing Hu, Luoning Gou, and Lu-Lu Chen. "Low-glycemic index diets as an intervention for diabetes: a systematic review and meta-analysis." *The American journal of clinical nutrition* 110, no. 4 (2019): 891-902. https://doi.org/10.1093/ajcn/nqz149

[3] Unwin, David, Ali Ahsan Khalid, Jen Unwin, Dominic Crocombe, Christine Delon, Kathy Martyn, Rajna Golubic, and Sumantra Ray. "Insights from a general practice service evaluation supporting a lower carbohydrate diet in patients with type 2 diabetes mellitus and prediabetes: a secondary analysis of routine clinic data including HbA1c, weight and prescribing over 6 years." *BMJ Nutrition, Prevention & Health* 3, no. 2 (2020): 285. https://doi.org/10.1136/bmjnph-2020-000072

[4] Jenkins, David JA, Cyril WC Kendall, Livia SA Augustin, Silvia Franceschi, Maryam Hamidi, Augustine Marchie, Alexandra L. Jenkins, and Mette Axelsen. "Glycemic index: overview of implications in health and disease." *The American journal of clinical nutrition* 76, no. 1 (2002): 266S-273S. https://doi.org/10.1093/ajcn/76.1.266S

[5] Saadati, Najmieh, Fatemeh Haidari, Mojgan Barati, Roshan Nikbakht, Golshan Mirmomeni, and Fakher Rahim. "The effect of low glycemic index diet on the reproductive and clinical profile in women with polycystic ovarian syndrome: A systematic review and meta-analysis." *Heliyon* 7, no. 11 (2021). https://doi.org/10.1016/j.heliyon.2021.e08338

[6] Marsh, Kate A., Katharine S. Steinbeck, Fiona S. Atkinson, Peter Petocz, and Jennie C. Brand-Miller. "Effect of a low glycemic index compared with a conventional healthy diet on polycystic ovary syndrome." *The American journal of clinical nutrition* 92, no. 1 (2010): 83-92. https://doi.org/10.3945/ajcn.2010.29261

[7] Livesey, Geoffrey, Richard Taylor, Helen F. Livesey, Anette E. Buyken, David JA Jenkins, Livia SA Augustin, John L. Sievenpiper *et al.*, "Dietary glycemic index and load and the risk of type 2 diabetes: a systematic review and updated meta-analyses of prospective cohort studies." *Nutrients* 11, no. 6 (2019): 1280. https://doi.org/10.3390/nu11061280

[8] Joint, F. A. O. "Carbohydrates in human nutrition." (1998).

[9] Brouns, F., I. Bjorck, K. N. Frayn, A. L. Gibbs, V. Lang, G. Slama, and T. M. S. Wolever. "Glycaemic index methodology." *Nutrition research reviews* 18, no. 1 (2005): 145-171. https://doi.org/10.1079/NRR2005100

[10] Henry, Christiani Jeyakumar, Rina Yu Chin Quek, Bhupinder Kaur, Sangeetha Shyam, and Harvinder Kaur Gilcharan Singh. "A glycaemic index compendium of non-western foods." *Nutrition & diabetes* 11, no. 1 (2021): 2. https://doi.org/10.1038/s41387-020-00145-w

[11] Osman, M. H., B. N. Yusof, and A. Ismail. "Glycaemic index and glycaemic load of foods and food products in Malaysia: A review." *International Food Research Journal* 28, no. 2 (2021): 217-229. https://doi.org/10.47836/ifrj.28.2.01

[12] Foster-Powell, Kaye, Susanna HA Holt, and Janette C. Brand-Miller. "International table of glycemic index and glycemic load values: 2002." *The American journal of clinical nutrition* 76, no. 1 (2002): 5-56. https://doi.org/10.1093/ajcn/76.1.5

[13] MacIntosh, Caroline G., Susanna HA Holt, and Jennie C. Brand-Miller. "The degree of fat saturation does not alter glycemic, insulinemic or satiety responses to a starchy staple in healthy men." *The Journal of nutrition* 133, no. 8 (2003): 2577-2580. https://doi.org/10.1093/jn/133.8.2577

[14] Amoah, Isaac, Carolyn Cairncross, Fabrice Merien, and Elaine Rush. "Glycaemic and appetite suppression effect of a vegetable-enriched bread." *Nutrients* 13, no. 12 (2021): 4277. https://doi.org/10.3390/nu13124277

[15] Arora, Kashika, Hana Ameur, Andrea Polo, Raffaella Di Cagno, Carlo Giuseppe Rizzello, and Marco Gobbetti. "Thirty years of knowledge on sourdough fermentation: A systematic review." *Trends in Food Science & Technology* 108 (2021): 71-83. https://doi.org/10.1016/j.tifs.2020.12.008

[16] Bajić, Filip, and Josip Job. "Data Extraction of Circular-Shaped and Grid-like Chart Images." *Journal of Imaging* 8, no. 5 (2022): 136. https://doi.org/10.3390/jimaging8050136

[17] Abràmoff, Michael D., Paulo J. Magalhães, and Sunanda J. Ram. "Image processing with ImageJ." *Biophotonics international* 11, no. 7 (2004): 36-42.

[18] Schneider, Caroline A., Wayne S. Rasband, and Kevin W. Eliceiri. "NIH Image to ImageJ: 25 years of image analysis." *Nature methods* 9, no. 7 (2012): 671-675. https://doi.org/10.1038/nmeth.2089

[19] Zhao, Ruixuan, Nan Li, Qiannan Liu, Wei Liu, Liang Zhang, and Honghai Hu. "Potato Flour, Oat Bran, and Inulin as Functional Ingredients in Gluten-Free Biscuits: Glycemic Index Reduction and Physicochemical Characterization Improvement." *Food and Bioprocess Technology* (2023): 1-12. https://doi.org/10.1007/s11947-023-03082-5

[20] Flint, Anne, Bente K. Møller, Anne Raben, Dorthe Pedersen, Inge Tetens, Jens J. Holst, and Arne Astrup. "The use of glycaemic index tables to predict glycaemic index of composite breakfast meals." *British Journal of Nutrition* 91, no. 6 (2004): 979-989. https://doi.org/10.1079/BJN20041124

[21] Eldakhakhny, Basmah Medhat, Aliaa Amr Alamoudi, Sarah Khalid Binmahfooz, Sarah Amr Alamoudi, Maha Essam Akshawi, and Jana Hisham Linjawi. "The Glycemic Index of Food-Should It Be Gender Specific?." *Current Research in Nutrition and Food Science Journal* 9, no. 1 (2021): 202-210. https://doi.org/10.12944/CRNFSJ.9.1.20

[22] Akinyede, Adedamola Iyioluwa, Adeola Bimbola Fehintola, and Timilehin David Oluwajuyitan. "Antioxidant activity and blood glucose reduction potential of Malabar Chestnut in Streptozotocin Induced Diabetic Rats." *Journal of Agriculture and Food Research* 8 (2022): 100299. https://doi.org/10.1016/j.jafr.2022.100299

[23] Chaturvedi, Neelam, Ayushi Agarwal, and Nishtha Raj. "Glycemic indices of Kale leaves based chakli and twisters on normal human subjects." *Int. J. Sci. Res* 6 (2017): 4.

[24] Almousa, Areej, Merciamma Thomas, Hisham Siddieg, Shiby Varghese, and Salahedeen Abusnana. "The glycemic index of traditional types of bread in UAE." *J. Nutr. Food Sci* 3 (2013). https://doi.org/10.4172/2155-9600.1000203

[25] Caballero-de la Peña, Katia, Laura Acevedo-Pacheco, Aidee I. Sánchez-Reséndiz, Cristina Chuck-Hernández, and Sergio O. Serna-Saldívar. "Protein quality and glycemic indexes of mango drinks fortified with a soybean/maize protein isolate with three levels of urease activity fed to weanling rats." *Food & Nutrition Research* 66 (2022). https://doi.org/10.29219/fnr.v66.8576

[26] Geetha, K., Geetha M. Yankanchi, Savita Hulamani, and Netravati Hiremath. "Glycemic index of millet based food mix and its effect on pre diabetic subjects." *Journal of food science and technology* 57 (2020): 2732-2738. https://doi.org/10.1007/s13197-020-04309-5

[27] Saculinggan, Mayette, and Emily Amor Balase. "Empirical power comparison of goodness of fit tests for normality in the presence of outliers." In *Journal of Physics: Conference Series*, vol. 435, no. 1, p. 012041. IOP Publishing, 2013. https://doi.org/10.1088/1742-6596/435/1/012041

[28] Sullivan, Gail M., and Anthony R. Artino Jr. "Analyzing and interpreting data from Likert-type scales." *Journal of graduate medical education* 5, no. 4 (2013): 541-542. https://doi.org/10.4300/JGME-5-4-18

[29] Lawrence, I., and Kuei Lin. "A concordance correlation coefficient to evaluate reproducibility." *Biometrics* (1989): 255-268. https://doi.org/10.2307/2532051

[30] Barnhart, Huiman X., Michael Haber, and Jingli Song. "Overall concordance correlation coefficient for evaluating agreement among multiple observers." *Biometrics* 58, no. 4 (2002): 1020-1027. https://doi.org/10.1111/j.0006-341X.2002.01020.x

[31] Bland, J. Martin, and DouglasG Altman. "Statistical methods for assessing agreement between two methods of clinical measurement." *The lancet* 327, no. 8476 (1986): 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8

[32] Akinjayeju, O., O. S. Ijarotimi, O. O. Awolu, and T. N. Fagbemi. "Nutritional composition, Glycaemic properties and anti-diabetic potentials of cereal-based soy-fortified flours for functional dough meal in diabetic induced rats." *Journal of Food Science and Nutrition Research* 3, no. 2 (2020): 102-120.

[33] Assefa, Meaza, Melaku Umeta, and Natesan Gnanaseknran. "Glycemic index of some traditional Ethiopian foods." *International Journal of Biochemistry Research & Review* 18, no. 4 (2017): 1-10. https://doi.org/10.9734/IJBCRR/2017/34471

[34] Blair, Robert M., E. C. Henley, and Aaron Tabor. "Soy foods have low glycemic and insulin response indices in normal weight subjects." *Nutrition journal* 5, no. 1 (2006): 1-10. https://doi.org/10.1186/1475-2891-5-35

[35] Bornet, F. R., Dominique Costagliola, S. W. Rizkalla, A. Blayo, A. M. Fontvieille, M. J. Haardt, M. Letanoux, G. Tchobroutsky, and G. Slama. "Insulinemic and glycemic indexes of six starch-rich foods taken alone and in a mixed meal by type 2 diabetics." *The American journal of clinical nutrition* 45, no. 3 (1987): 588-595. https://doi.org/10.1093/ajcn/45.3.588

[36] Di Cairano, Maria, Fideline Laure Tchuenbou-Magaia, Nicola Condelli, Nazarena Cela, Constance Chizoma Ojo, Iza Radecka, Simon Dunmore, and Fernanda Galgano. "Glycaemic Index of Gluten-Free Biscuits with Resistant Starch and Sucrose Replacers: An In Vivo and In Vitro Comparative Study." *Foods* 11, no. 20 (2022): 3253. https://doi.org/10.3390/foods11203253

[37] Jimoh, A. K., S. O. Adediran, S. A. Adebisi, E. K. Oghagbon, and A. B. Okesina. "The mode of preparation of white yam (dioscorea rotunda) affects the glycaemic response in type II diabetic Nigerian subjects." *Proceedings of the Nutrition Society* 72, no. OCE1 (2013): E34. https://doi.org/10.1017/S0029665113000360

[38] Luke, David R., Karen Ka Yan Lee, Carl W. Rausch, and Camille Cheng. "Phase 1 Study of the Pharmacology of BTI320 Before High-Glycemic Meals." *Clinical pharmacology in drug development* 8, no. 3 (2019): 395-403. https://doi.org/10.1002/cpdd.583

[39] Nagaraju, Raju, Prasanthi Prabhakaran Sobhana, Deepika Thappatla, Srinivas Epparapalli, Bhaskarachary Kandlakunta, and Damayanti Korrapati. "Glycemic index and sensory evaluation of whole grain based multigrain Indian breads (Rotis)." *Preventive Nutrition and Food Science* 25, no. 2 (2020): 194. https://doi.org/10.3746/pnf.2020.25.2.194

[40] Olugbuyi, Ayo Oluwadunsin, Sunday Abiodun Malomo, Oluwole Steve Ijarotimi, and Tayo Nathaniel Fagbemi. "Amino acids profile, glyceamic index/load, in-vitro antioxidant and sensory attributes of optimized dough meal from the blends of plantain, Soycake and rice-bran flours." *Journal of Culinary Science & Technology* 21, no. 5 (2023): 795-817. https://doi.org/10.1080/15428052.2021.2016530

[41] Oluwajuyitan, Timilehin David, and Oluwole Steve Ijarotimi. "Nutritional, antioxidant, glycaemic index and Antihyperglycaemic properties of improved traditional plantain-based (Musa AAB) dough meal enriched with tigernut (Cyperus esculentus) and defatted soybean (Glycine max) flour for diabetic patients." *Heliyon* 5, no. 4 (2019). https://doi.org/10.1016/j.heliyon.2019.e01504

[42] Origbemisoye, Babawande A., and Beatrice O. Ifesan. "Nutritional, phytochemical, glycaemic index, antihyperglycaemic properties and quality attribute of Wheat-ackee arils cookies for diabetic patients." *Journal of Food Measurement and Characterization* 15 (2021): 144-154. https://doi.org/10.1007/s11694-020-00623-w

[43] Raczkowska, Ewa, Karolina Łoźna, Maciej Bienkiewicz, Karolina Jurczok, and Monika Bronkowska. "Relation between the recipe of yeast dough dishes and their glycaemic indices and loads." *Foods* 8, no. 9 (2019): 377. https://doi.org/10.3390/foods8090377

[44] McBride, G. B. "A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient." *NIWA client report: HAM2005-062* 45 (2005): 307-310.

[45] Altman, Douglas G., and J. Martin Bland. "Measurement in medicine: the analysis of method comparison studies." *Journal of the Royal Statistical Society Series D: The Statistician* 32, no. 3 (1983): 307-317. https://doi.org/10.2307/2987937

[46] Gheibi, Sevda, Alireza Mahmoodzadeh, Khosrow Kashfi, Sajad Jeddi, and Asghar Ghasemi. "Data extraction from graphs using Adobe Photoshop: applications for meta-analyses." *International Journal of Endocrinology and Metabolism* 17, no. 4 (2019). https://doi.org/10.5812/ijem.95216

[47] Rane, Chinmayee, Seshasayee Mahadevan Subramanya, Devi Sandeep Endluri, Jian Wu, and C. Lee Giles. "Chartreader: Automatic parsing of bar-plots." In *2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI)*, pp. 318-325. IEEE, 2021. https://doi.org/10.1109/IRI51335.2021.00050

[48] Al-Zaidy, Rabah A., and C. Lee Giles. "Automatic extraction of data from bar charts." In *Proceedings of the 8th international conference on knowledge capture*, pp. 1-4. 2015. https://doi.org/10.1145/2815833.2816956

[49] Mishra, Prerna, Santosh Kumar, and Mithilesh Kumar Chaube. "Evaginating scientific charts: Recovering direct and derived information encodings from chart images." *Journal of Visualization* (2022): 1-17.

[50] Atkinson, Fiona S., Jennie C. Brand-Miller, Kaye Foster-Powell, Anette E. Buyken, and Janina Goletzke. "International tables of glycemic index and glycemic load values 2021: a systematic review." *The American journal of clinical nutrition* 114, no. 5 (2021): 1625-1632. https://doi.org/10.1093/ajcn/nqab233

[51] Aston, Louise M. "Glycaemic index and metabolic disease risk." *Proceedings of the Nutrition Society* 65, no. 1 (2006): 125-134. https://doi.org/10.1079/PNS2005485

[52] Chiavaroli, Laura, Danielle Lee, Amna Ahmed, Annette Cheung, Tauseef A. Khan, Sonia Blanco, Arash Mirrahimi *et al.,* "Effect of low glycaemic index or load dietary patterns on glycaemic control and cardiometabolic risk factors in diabetes: systematic review and meta-analysis of randomised controlled trials." *Bmj* 374 (2021). https://doi.org/10.1136/bmj.n1651

[53] Luo, Junyu, Zekun Li, Jinpeng Wang, and Chin-Yew Lin. "Chartocr: Data extraction from charts images via a deep hybrid framework." In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 1917-1925. 2021. https://doi.org/10.1109/WACV48630.2021.00196

[54] Moeyaert, Mariola, Daniel Maggin, and Jay Verkuilen. "Reliability, validity, and usability of data extraction programs for single-case research designs." *Behavior Modification* 40, no. 6 (2016): 874-900. https://doi.org/10.1177/0145445516645763

[55] Drevon, Daniel, Sophie R. Fursa, and Allura L. Malcolm. "Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data." *Behavior modification* 41, no. 2 (2017): 323-339. https://doi.org/10.1177/0145445516673998

[56] Liu, Jin Long, Xue Jun Zhang, Zhao Zhang, An Hong Zhang, Wei Wang, and Jia Hong Dong. "Meta-analysis: prognostic value of survivin in patients with hepatocellular carcinoma." *PLoS One* 8, no. 12 (2013): e83350. https://doi.org/10.1371/journal.pone.0083350