OPEN ACCESS

Presentation of clinical laboratory results: an experimental comparison of four visualization techniques

Torbjørn Torsvik,[1] Børge Lillebo,[1] Gustav Mikkelsen[2]

[1]Norwegian EPR Research Centre, Department of Neuroscience, The Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway
[2]Department of Laboratory Medicine, Children's and Women's Health, The Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway

Correspondence to
Børge Lillebo, Faculty of Medicine/Department of Neuroscience, Norwegian EPR Research Centre, Norwegian University of Science and Technology, Pb 8905, MTFS, Trondheim 7491, Norway; borge.lillebo@ntnu.no

ABSTRACT

Objective To evaluate how clinical chemistry test results were assessed by volunteers when presented with four different visualization techniques.

Materials and methods A total of 20 medical students reviewed quantitative test results from 4 patients using 4 different visualization techniques in a balanced, crossover experiment. The laboratory data represented relevant patient categories, including simple, emergency, chronic and complex patients. Participants answered questions about trend, overall levels and covariation of test results. Answers and assessment times were recorded and participants were interviewed on their preference of visualization technique.

Results Assessment of results and the time used varied between visualization techniques. With sparklines and relative multigraphs participants made faster assessments. With relative multigraphs participants identified more covarying test results. With absolute multigraphs participants found more trends. With sparklines participants more often assessed laboratory results to be within reference ranges. Different visualization techniques were preferred for the four different patient categories. No participant preferred absolute multigraphs for any patient.

Discussion Assessments of clinical chemistry test results were influenced by how they were presented. Importantly though, this association depended on the complexity of the result sets, and none of the visualization techniques appeared to be ideal in all settings.

Conclusions Sparklines and relative multigraphs seem to be favorable techniques for presenting complex long-term clinical chemistry test results, while tables seem to suffice for simpler result sets.

BACKGROUND AND SIGNIFICANCE

The importance of laboratory test results in clinical work is unquestionable. In hospital settings, laboratory test use seems to be increasing considerably,[1] and in primary care, physicians may receive as many as 1000 test results each week.[2] Physicians have to stay aware of new results, comprehend the results and ensure proper follow-up based on assessment of single and multiple values and systematic changes over time. Studies have shown that these tasks are not straightforward. Physicians may be unaware of abnormal test results, and abnormal results may be left unrecognized without proper follow-up.[3–5]

A single clinical laboratory test result may consist of a numeric value—representing the concentration of a substance in for example, the patient's blood—

accompanied by the name of the test, the unit of measurement, the date of the sampling and a reference range. The reference range is commonly defined as the 95% central range of values observed in healthy individuals. Clinicians may compare individual test results with the reference range for the test, in order to establish whether the result is high or low compared to healthy individuals.

The laboratory report is a vital link between the laboratory and the physician, and the presentation format can have major impact on the clinical action taken.[6] Traditionally, laboratory results have been presented as tables. This has probably been related to use of paper based patient records, and the simplicity of adding new entries of laboratory results into a table. However, electronic health information systems permit visualizing these results in alternative ways. One study showed that laboratory data presented with one particular line graph visualization—'sparklines'—were assessed faster than when presented in a conventional table,[7] while non-clinical studies have come to the opposite conclusion.[8][9] A problem with comparing studies of visualization techniques is that there are numerous ways to present laboratory results.[10][11] Additionally, clinical contexts differ, and it is not certain that one technique fits all clinical situations.

OBJECTIVE

In this study we evaluated how four different visualization techniques—three line graphs and one table—performed when presenting numerical clinical chemistry test results from four patients, each representing a distinct patient category: the emergency patient, the chronic patient, the simple patient, and the complex patient. We focused on how trends, overall levels and covariation were assessed with different visualization techniques, including assessment times. In addition we evaluated subjective user preferences with respect to the four techniques.

Two of the visualization techniques, the table and the absolute multigraph, were based on solutions implemented in hospital and primary care systems in our region. The third visualization technique—sparklines—has been described and studied by others and was thus highly relevant for comparison with the other techniques.[7][11][12] With the fourth technique—the relative multigraph—we tried to solve some of the problems with simultaneous visualization of multiple tests with the absolute multigraph. This was somewhat inspired by the unit-independent technique,[10] but rather than scaling results by test SD and using a

logarithmic time axis, the relative multigraph had a (partially) logarithmic value axis and a linear time axis.

METHODS AND MATERIALS
Study design
Deidentified clinical chemistry test results from four patients were presented to each participant using the four visualization techniques in a balanced, crossover experiment. The study was conducted during May 2011 at The Norwegian EPR Research Centre at the Norwegian University of Science and Technology (NTNU).

Participants
A total of 20 medical students (9 women) at the NTNU were recruited through mailing lists, posters and direct contact. Participation was stimulated by a gift coupon that would be given to one of the participants. Their mean age was 25.3 years, and their mean length of studying medicine was 3.4 years (range 1–5 years). The medical faculty at NTNU has an integrated curriculum that involves problem-based learning and student–patient and student–physician sessions from the first year of studies. Thus, all included students were expected to have general knowledge about assessment of laboratory test results.

Visualization techniques
The four visualization techniques that were studied are illustrated in figure 1.

In the table, the names of the laboratory tests together with their respective reference ranges were listed as separate rows in the first column. Subsequent columns listed test results for individual samples in reverse chronological order (ie, most recent samples to the left). The column headers displayed the date and time of sample collection. Values outside the reference range were colored red and labeled either 'H' (high) or 'L' (low). When results from many samples were presented in the same table (the chronic and the complex patient cases), the user had to scroll horizontally to see all results within the boundaries of the display.

The sparklines visualization displayed laboratory data as miniature line graphs in separate miniature reference systems with vertical axes adapted to the range of results, and horizontal axes representing a common time frame. This technique has been referred to as 'word-sized graphics'.[12] Each sparkline included a line representing the results and a shaded field representing the reference range for that particular test. A label above the sparkline stated the name of the test.

The absolute multigraph also visualized laboratory data as line graphs with reference range fields for each line, but unlike sparklines all lines and reference range fields were plotted within the same reference system with the horizontal axis representing time and the vertical axis representing the total range of numerical values in the data. Both axes were linear. This technique had some obvious problems. For instance, a serious drop in hemoglobin levels (reference ranges 13.4–17.0) would hardly be visible when plotted within a reference system with a vertical axis from 0 to 500 (eg, together with platelet counts). This problem could be circumvented through interaction with the visualization by displaying only those tests that were of interest, since the vertical axis synchronously adjusted to fit the values of the selected tests only. This interaction was performed by clicking on the name of the tests in the legend below the visualization, which had color coding to facilitate identification of the tests in the plot. Problems with visualization of multiple tests with different ranges plotted together in a common reference system has been discussed elsewhere.[7]

Finally, we constructed a relative multigraph. Like the absolute multigraph it visualized laboratory data as separate line graphs within a common coordinate system, and it had a similar interactive and color-coded legend. But unlike the absolute multigraph all test values were transformed according to the width of each test's reference ranges, in order to fit a common scale and reference range on the y axis. In addition, the y axis was linear within the reference range and logarithmic outside.

No numerical values were visible in any of the line graph visualizations, and all line graphs were plotted in the opposite chronological order to that of the table (ie, line graphs had
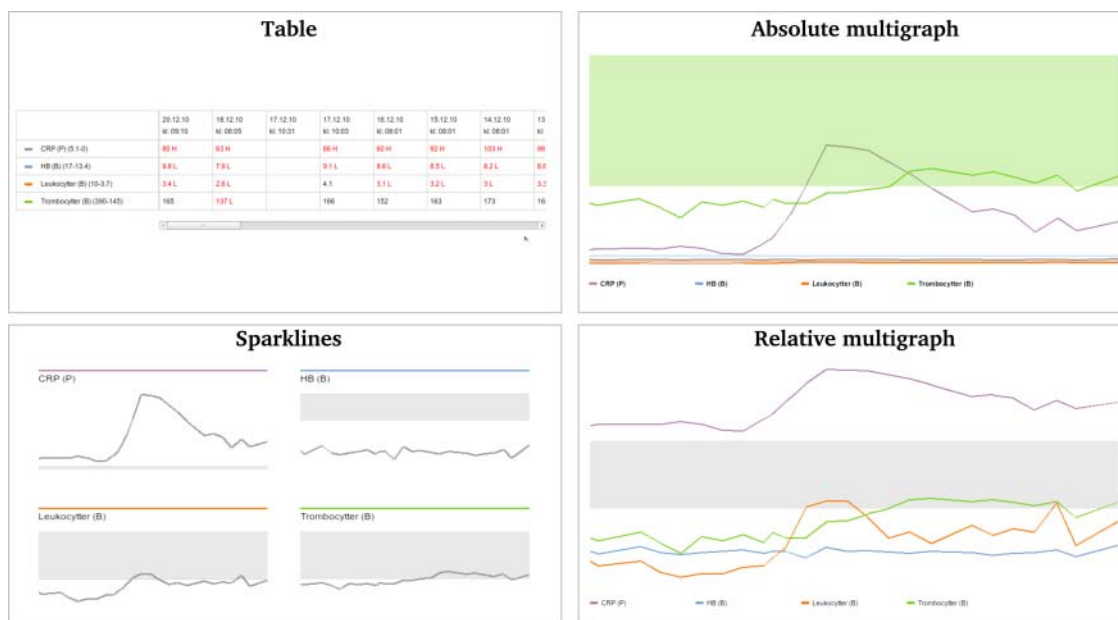

Figure 1 Four visualizations of the same laboratory data (the chronic patient case).

most recent results to the right). We chose to do this based on our experience with presentation formats of laboratory reports in existing patient record systems.

Patient cases

Each visualization was applied to laboratory data from four patients (table 1). The laboratory data were chosen to reflect different patient categories for which laboratory test results would have to be interpreted. No other information pertaining to the cases were given.

Procedure

Before the experiment each participant was informed about the project and the four visualization techniques. Participants practiced approximately 10 min on how to interact with the visualizations and how to submit their answers using keyboard and mouse. They were told to answer correctly and as fast as they could. The tests were performed with a desktop computer. The software was programmed in php and JavaScript using a MySQL database. Visualizations were shown sequentially in a 950×450 px area on a 1920×1080 px monitor. Participants were told that they would see laboratory data from many patients with varying visualization techniques. They were not told that there were only four different sets of laboratory data each visualized with four different techniques and presented in a predefined mixed order making each participant his or her own control. The presentation order of the visualizations was changed between each participant to avoid ordering effects (relative-sparklines-table-absolute, sparklines-table-absolute-relative, table-absolute-relative-sparklines or absolute-relative-sparklines-table). The order of cases was the same for all participants (chronic-complex-emergency-simple-complex-emergency-simple-chronic-emergency-simple-chronic-complex-simple-chronic-complex-emergency).

After the experiment participants were informed that there were in fact only four different cases, and they were interviewed on their preference among the four visualization

Table 1 Overview of laboratory data that were presented in each patient case

Patient case	No. of results	No. of samples	No. of tests	Tests
Simple	10	3	4	P: alanine aminotransferase, C reactive protein, creatinine, potassium
Emergency	35	3	15	P: alanine aminotransferase, albumin, alkaline phosphatase, amylase, bilirubin, C reactive protein, creatinine, γ-glutamyl transferase, glucose, PT-INR, potassium, sodium
				B: hemoglobin, platelet count, white blood cell count
Chronic	101	26	4	P: C reactive protein
				B: hemoglobin, platelet count, white blood cell count
Complex	233	23	15	P: alanine aminotransferase, C reactive protein, creatinine, γ-glutamyl transferase, glucose, magnesium, potassium, sodium
				B: basophil granulocyte count, hemoglobin, neutrophil granulocyte count, platelet count, white blood cell count
				VB: bicarbonate, carbon dioxide partial pressure

Not all tests were run for each sample.
B, whole blood; P, plasma; PT-INR, prothrombin time/ international normalized ratio; VB, venous whole blood.

Table 2 Questions that the participants had to answer

Category	Question	Answer
Trend	Do you consider the results of 'test X' to have increased/decreased significantly during the period?	Increased Decreased Neither
Overall levels	Overall, do you consider the results of 'test X' to be above/below the reference ranges?	Above Below Neither
Covariation	When you consider all tests for this patient, can you see any covariation between any of the results?	Free text

techniques for each of the four cases. Each experiment lasted approximately 1 h.

Outcome measures

For each combination of case and visualization technique the participants had to answer three questions (table 2). Answers were automatically recorded in a database together with the time the participant spent from the question appeared on the screen until a submit button was clicked. For assessments of trends and overall levels, a mean assessment time per test was calculated by dividing the recorded time with the number of tests covered by each question (only 5 of the 15 tests had to be assessed for the emergency and complex cases as opposed to all 4 for the simple and chronic cases).

After all experiments were completed, the free text comments on covariation were coded independently by two of the authors blinded for what visualization that triggered the comment. We only considered covariation comments for the complex patient case (many tests and many samples). Our definition of covariation was synchronous changes of two or more tests (eg, 'C reactive protein (CRP) and leukocytes increase at the same time'). The coders gave each test mentioned in a covariation comment 1 point.

Statistical analysis

Because there were no valid criteria for how the laboratory results should be assessed with respect to trends and overall levels, our focus was on pairwise analyses of agreement (Cohen's κ) and disagreement (McNemar's test) between visualization techniques—that is, intervisualization agreement and disagreement (comparable to inter-rater agreement in reliability studies). That is, to what extent assessments of identical laboratory data were identical or consistently different between visualization techniques.

Assessment times for trends and overall levels were analyzed using mixed model analysis with participant as a random effect and visualization technique, patient case and repeated exposure as fixed effects. Repeated exposure referred to repetition of visualization technique and patient case due to the balanced, crossover design of this study.

Differences in covariation scores between visualization techniques were tested for statistical significance with the nonparametric Friedman test.

Finally, preferred visualization techniques for each patient case were analyzed with the exact multinomial test, presuming a uniform distribution between visualization techniques. Statistical analyses were performed with IBM SPSS Statistics (V.19; SPSS, Chicago, Illinois, USA), R 2.01 software (www.r-project.org) and SAS V.9.3 (SAS Institute Inc., Cary, North Carolina, USA).

Research and applications
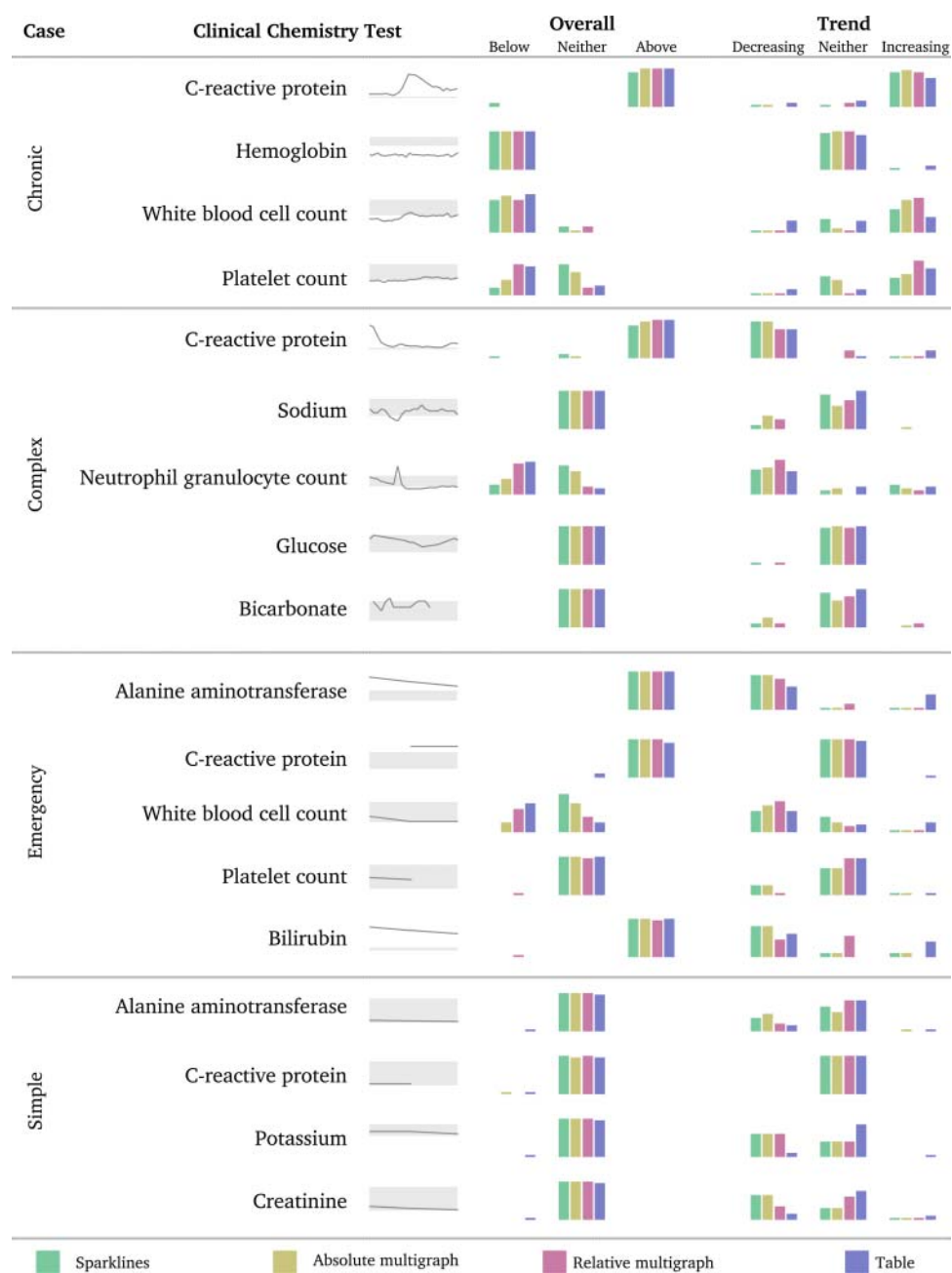
Figure 2 Laboratory test results for the four different patients are displayed as sparklines. Relative distributions of answers to questions about trends and overall levels for these tests are displayed as vertical bars for the four visualization techniques.

Ethics

The Data Protection Official for Research for Norwegian universities (NSD) was consulted before the study and concluded that no further approval was required since only anonymous data were collected.
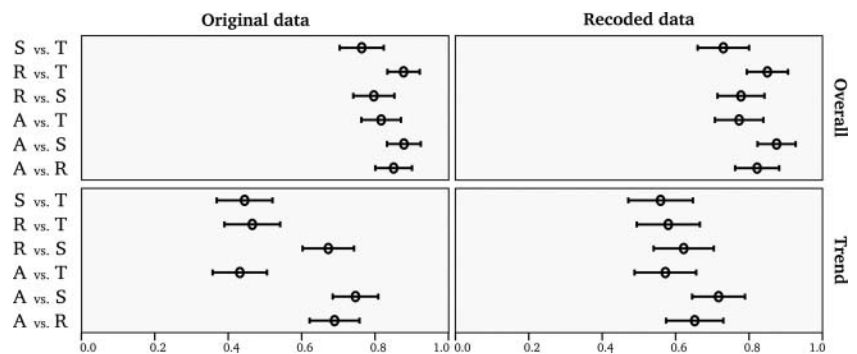
RESULTS

Assessment of trends and overall levels

In total the 20 participants made 2880 assessments of trends and overall features of laboratory test results (figures 2 and 3). In general, agreement between visualization techniques was higher for overall level assessments than trend assessments. Pairs consisting of the table and any other line graph visualization had statistically significant poorer agreement with respect to assessment of trend compared to pairs of two line graph visualizations (CI not overlapping in figure 3).

Inspection of the data indicated that some participants had wrongfully assessed the time course in the table as going from left to right (figure 2), causing lower agreement between table and the line graphs for the trend assessment (figure 3). For instance, eight participants assessed the bilirubin levels of the emergency case presented with the table as an increasing trend although the values clearly demonstrated a decreasing trend (in reverse chronological order the bilirubin levels were 161, 195, and 231). Similar flaws were observed for other assessments as well (figure 4). However, the trends of some of these tests had increasing and decreasing segments, thus complicating any certain conclusions as to whether the participant misinterpreted the time course or merely assessed the trend differently. Additionally, the trend features of tests presented with the other visualization techniques were sometimes wrongfully assessed as well (figure 4).

Figure 3 Pairwise Cohen's κ with CI between all six possible pairs of the four visualization techniques (intervisualization agreement) regarding questions about trend and overall levels. Original data (trinomial) and recoded data (binomial) are presented (recoding is further explained in the text). A, absolute multigraph; R, relative multigraph; S, sparklines; T, table.

There was no apparent pattern in the results that allowed us to adjust for this misconception, hence we chose to recode our data from a trinomial assessment (decreasing, increasing and neither) to a binomial (any trend vs no trend). In this way, our data was not affected by participants misinterpreting the time. Correspondingly, we also recoded the overall assessment data from trinomial (above, below and neither) to binomial (within vs beyond reference ranges) (figure 3).

Whenever laboratory results were assessed differently between visualization techniques (disagreement), laboratory data presented with absolute multigraph were consistently more likely to be assessed as a trend (increasing or decreasing) compared to the other techniques (McNemar test: table p<0.001; sparklines p=0.005; and relative multigraph p=0.002). Additionally the absolute multigraph was less likely to be assessed as being beyond reference ranges compared to table (p<0.001) and relative multigraph (p<0.001). Laboratory results presented with sparklines were consistently less likely to be assessed as beyond reference ranges compared to the other visualization techniques (absolute multigraph p=0.001; relative multigraph p<0.001; table p<0.001). There were no statistically significant differences between relative multigraph and table with respect to assessments of overall levels (p=0.248) and trend (p=0.106), nor was there any significant differences in trend assessments between relative multigraph and sparklines (p=0.716), or between sparklines and table (p=0.043, Bonferroni correction requires p<0.008).

Assessment times

Assessment times differed between visualization techniques as well as between patient cases, questions and repeated exposure. The shortest assessment times were achieved with sparklines and relative multigraphs presenting the laboratory results for the emergency and simple cases as the third or fourth exposure (figure 4). The experiment was not designed to identify differences between question types as overall levels were always assessed right before trends—favoring trend assessments.

By mixed model analysis we found significant interaction between visualization technique, patient case and repeated exposure (p<0.001), indicating that the visualization techniques had different effects on assessment time based on which case they presented, and the degree of repeated exposure to that case and visualization technique (ie, a differentiated learning effect). The association between assessment time and visualization technique was statistically significant for the chronic (p<0.001) and complex (p<0.009) cases, but not for the simple (0.082<p<0.713) and emergency (0.145<p<0.742) cases. This effect was consistent through all repeated exposures.

Due to small sample sizes when broken down into all possible combinations of visualization technique, patient case and repeated exposure—and because analyzing each exposure and case combination separately would break the within-subjects, repeated measures design—we did not do any further post-hoc statistical analyses. However, visual inspection of the data demonstrates that the most evident differences in assessment times are between the table and the three other visualization techniques for the chronic and the complex patient cases (figure 4). The figure also indicates that variation in assessment times decreased through repeated exposures. Sparklines and relative multigraph performed quite well through all repetitions, the table performed poorly, and the absolute multigraph somewhere in between.

Assessment of covariation

The agreement between the two investigators performing the coding of free text covariation comments was good (Cohen's κ 0.91) indicating valid interpretation of free text comments about covariations. The relative multigraphs generated the highest covariation score (table 3). The differences in covariation scores were statistically significant (Friedman test, $\chi^2(3)=10.853$, p=0.013). Post-hoc analyses with Wilcoxon
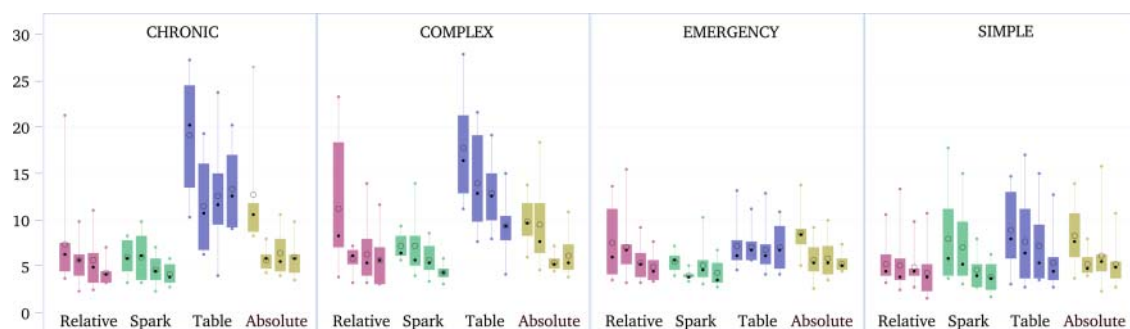
Figure 4 Boxplots of assessment times per patient case, visualization technique and repetition (increasing repetition from left to right among adjacent boxplots with identical color). Bars: 1st to 3rd quartile. Whiskers: minimum to maximum. Circle: mean. Dot: median.

Table 3 Number of tests commented as covarying with each other

Output	Covariation score
Table	18
Absolute multigraph	31
Sparklines	38
Relative multigraph	47

signed-rank tests for all six possible pairs of visualization techniques demonstrated statistically significant differences only between table and relative multigraph (p=0.003) and table and sparklines (p=0.013).

User preferences

The table was preferred by most participants for the simple patient case and was also one of the most preferred techniques for visualizing the emergency patient case. Because values outside the reference interval were written in red letters, they were easy to spot. Many participants felt that graphical visualization models in general lost their usefulness with low number of samples.

The relative multigraph was the most preferred technique for visualizing the chronic patient case. Participants explained that the relative multigraph provided the best overview when few tests were to be presented, and that the common timeline facilitated perception of covariations. Many participants also preferred the relative multigraph for the simple patient because they could immediately see that all test results were within reference ranges when no lines went beyond the fixed reference area. With more tests, many participants felt that the relative multigraph became too clogged up with lines and therefore preferred sparklines. Sparklines were characterized as easy to understand and as giving a good overview of laboratory results irrespective of patient case. No participants preferred the absolute multigraph for any patient case.

Multinomial exact tests found statistically significant deviations from uniform distribution of preference for each patient case (table 4).

Although participants assessed only four different sets of laboratory results, it was very difficult for them to know this for certain. A majority of the participants said that they suspected that some of the 16 visualizations presented the same laboratory results, but they did not believe that this affected their assessments.

DISCUSSION

This study demonstrates that there are differences between visualization techniques with respect to how laboratory results are assessed and how fast the assessments are made. Additionally, the characteristics of the laboratory data presented with these techniques affected user preference and assessment times. To our knowledge, this is the first time different patient categories have been included in an evaluation of visualization techniques for presentation of clinical laboratory results, and the first time several line graph techniques have been compared with each other.

For small sets of laboratory data, a table seems to be sufficient and preferable—especially for few samples as with a new emergency patient. However, whenever repeated tests (many samples) have to be assessed—for example, monitoring glucose or creatinine levels in chronically ill patients—line graph visualizations are assessed approximately twice as fast and are more preferred than a table. We observed only moderate variation in assessment times between different line graph techniques, but the relative multigraph and sparklines provided faster assessments than the absolute multigraph.

In general, agreement between visualization techniques was good for assessments of trend or overall levels. However, whenever assessments disagreed, the techniques demonstrated different propensity for how they were assessed. Laboratory results presented with the absolute multigraph were more often interpreted as decreasing or increasing trends, and results presented with sparklines were less often interpreted as beyond reference ranges. These differences are not surprising. The absolute multigraph was not suitable for presenting several tests simultaneously on a common value axis. However, through interaction it was possible to visualize tests one by one. In that way, each test was presented using maximal screen estate. Thus, even small increasing or decreasing trends would be more noticeable compared to the other techniques. With the table, red color on laboratory results that were beyond reference ranges gave an immediate impression of overall levels. Even though results would be barely outside of reference range, the red color was striking to the eye. This contrasts the line graph techniques, especially sparklines, which provided the lowest resolution per visualized test among the line graph techniques. A small deviation from the reference range would not be easily spotted with sparklines since the line graph would be located on the edge corresponding to the reference range. No similarly consistent features were related to overall levels and trend assessments made with the relative multigraph, but it was the technique with which the participants most frequently indicated covarying results. An explanation for this could be that the relative multigraph presented the results with relation to a common timeline and a common reference range.

It seems very likely that some participants misinterpreted the time course of the table and assessed decreasing trends as increasing and vice versa. Such misinterpretations can obviously affect how patients are managed and should therefore be given much attention. However, in a clinical setting the laboratory results have to be combined with other clinical information and any pretest expectations, clinicians are able to actively choose the visualization technique they want based on what question they need to answer, and finally clinicians are probably more accustomed to the systems they are using. Thus, we think that this kind of error is less likely to occur in a clinical setting, yet this is a subject for further study.

One of the strengths with this study was that we included clinical chemistry test results from patients that represented different clinical problems, rather than presenting test results from similar patients. Additionally, we used a multimethod evaluation, including quantitative and qualitative approaches. However, this study also has limitations in that we did not ask any questions related to single numeric values, nor did we require the participants to combine laboratory data with clinical information in order to make more complex medical decisions on diagnosis, prognosis or therapy. Thus, the results should not be uncritically generalized to clinical settings. The

Table 4 Participants' preferred visualization technique for each case

Patient case	n	Table	Relative multigraph	Sparklines	Absolute multigraph	p Value
Simple patient	20	10	7	3	0	0.004
Emergency patient	20	9	2	9	0	0.001
Chronic patient	20	1	14	5	0	<0.001
Complex patient	20	2	2	16	0	<0.001

use of medical students as participants could also be regarded a limitation. However, the assessment they had to do did not require deep medical knowledge.

Bauer *et al* performed a similar experiment with a table and a sparkline visualization technique.[7] Their main results correspond with our results, namely that humans assess line graphs faster than tables and that interpretation of laboratory results may vary between visualization techniques. As we did, they also found that values slightly beyond reference ranges were more often identified with a table compared to sparklines. Similar experiments with graphical representations of numerical data from monitoring anesthetized patients have found that graphical displays can improve presentation of medical information.[13–15]

Bauer did not find any significant effect of repeated exposure to the cases, while in our results the learning effects of repeated exposure were statistically significant. This difference could be caused by the variations in experimental design between the studies. In the experiment by Bauer *et al*, 12 physicians interpreted 11–13 tests pertaining to each of 4 rather similar cases (ie, pediatric intensive care unit patients with identical test sets) visualized with 2 techniques (ie, 2 exposures) and submitting their answers through talk aloud technique. In our experiment 20 medical students interpreted 4–5 tests pertaining to each of 4 different cases visualized with four techniques (ie, 4 exposures) and submitting their answers in a computerized form. Our data does not provide any further insight into why our participants made faster assessments towards the end of the experiment. Possible explanations could be recognition of cases, familiarity with visualization techniques or merely improved mastering of the experimental situation.

As our results demonstrated, the characteristics of the data that were visualized had significant effects on how it was assessed. This makes it difficult to compare the results from different studies even though the visualization techniques are identical. Perhaps standards should be developed (standard laboratory data, standard patient cases) for how visualization techniques for laboratory—and even clinical—data should be experimentally evaluated to ensure sufficient methodological rigor?

Our results are not clear on what is the optimal visualization technique for laboratory data, rather they demonstrate advantages and disadvantages with different techniques. Before making more specific recommendations we would like to encourage studies with more complex questions and gold standards for comparison. Nevertheless, as Bauer *et al* and Tufte have shown, sparklines are easy to integrate in composite visualizations of tables and line graphs.[7 12] Additionally they consume little screen estate. On the other hand, a relative multigraph can more easily be integrated with a timeline oriented patient record, facilitating covariation analysis of laboratory data with other clinical information.[16] More research on such integrated views of laboratory data and non-laboratory clinical data should be performed in order to optimize clinical data presentation techniques.

This is the first time the relative multigraph is included as a visualization technique in an experiment with authentic laboratory data, and we are not aware of any clinical information system that presents laboratory results as a relative multigraph. A similar technique we have found described in literature is the unit-independent technique.[10] This technique has SD units on the value axis and a logarithmic time axis. A logarithmic time scale provides a long-term overview together with a more detailed presentation of recent results, but comparing time intervals may be more difficult than with linear time scales. Moreover, some problems are common to both of these techniques. One problem is presenting many tests together which may result in a clutter of lines that can be hard to separate from each other. Another problem is understanding the absolute values of a test by looking at the position on the y axis. These issues call for more research.

CONCLUSIONS

This study demonstrated that different techniques for visualizing and presenting numeric laboratory results influenced on how the results were assessed. For simple and acute patient problems with short time spans and few blood samples, a table seemed to suffice, but for more complex patient problems with long-term monitoring a relative multigraph or sparklines seemed favorable. More development has to be undertaken to improve these techniques and integrate them with other clinical non-laboratory information.

REFERENCES

1. **Mindemark M,** Larsson A. Longitudinal trends in laboratory test utilization at a large tertiary care university hospital in Sweden. *Upsala J Med Sci* 2011;**116**:34–8.
2. **Poon EG,** Wang SJ, Gandhi TK, *et al*. Design and implementation of a comprehensive outpatient Results Manager. *J Biomed Inform* 2003;**36**:80–91.
3. **Roy CL,** Poon EG, Karson AS, *et al*. Patient safety concerns arising from test results that return after hospital discharge. *Ann Intern Med* 2005;**143**:121–8.
4. **Edelman D.** Outpatient diagnostic errors: unrecognized hyperglycemia. *Eff Clin Pract* 2002;**5**:11–6.
5. **Callen J,** Georgiou A, Li J, *et al*. The safety implications of missed test results for hospitalised patients: a systematic review. *BMJ Qual Saf* 2011;**20**:194–9.
6. **Mayer M,** Wilkinson I, Heikkinen R, *et al*. Improved laboratory test selection and enhanced perception of test results as tools for cost-effective medicine. *Clin Chem Lab Med* 1998;**36**:683–90.
7. **Bauer DT,** Guerlain S, Brown PJ. The design and evaluation of a graphical display for laboratory data. *J Am Med Inform Assoc* 2010;**17**:416–24.
8. **Meyer J.** Performance with tables and graphs: effects of training and a Visual Search Model. *Ergonomics* 2000;**43**:1840–65.
9. **Lohse GL.** A cognitive model for understanding graphical perception. *HumComp Interact* 1993;**8**:353–88.
10. **Mayer M,** Chou D, Eytan T. Unit-independent reporting of laboratory test results. *Clin Chem Lab Med* 2001;**39**:50–2.
11. **Powsner SM,** Tufte ER. Graphical summary of patient status. *Lancet* 1994;**344**:386–9.
12. **Tufte E.** *Beautiful evidence*. Cheshire, CT: Graphics Press LLC, 2006:213.
13. **Wachter SB,** Johnson K, Albert R, *et al*. The Evaluation of a Pulmonary Display to Detect Adverse Respiratory Events Using High Resolution Human Simulator. *J Am Med Inform Assoc* 2006;**13**:635–42.
14. **Agutter J,** Drews F, Syroid N, *et al*. Evaluation of graphic cardiovascular display in a high-fidelity simulator. *Anesth Analg* 2003;**97**:1403–13.
15. **Gurushanthaiah K,** Weinger M, Englund C. Visual display format affects the ability of anesthesiologists to detect acute physiologic changes: a laboratory study employing a clinical display simulator. *Anesthesiology* 1995;**83**:1184–93.
16. **Plaisant C,** Mushlin R, Snyder A, *et al*. LifeLines: using visualization to enhance navigation and analysis of patient records. *Proc AMIA Symp* 1998:76–80.

Presentation of clinical laboratory results: an experimental comparison of four visualization techniques

Torbjørn Torsvik, Børge Lillebo and Gustav Mikkelsen

Updated information and services can be found at:

http://jamia.bmj.com/content/20/2/325.full.html

These include:

References	This article cites 14 articles, 3 of which can be accessed free at: **http://jamia.bmj.com/content/20/2/325.full.html#ref-list-1**
Open Access	This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: http://creativecommons.org/licenses/by-nc/3.0/ and http://creativecommons.org/licenses/by-nc/3.0/legalcode
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections	Articles on similar topics can be found in the following collections Open access (162 articles)

Notes