Usability and Suitability of the Omics-Integrating Analysis Platform tranSMART for Translational Research and Education

J. Christoph[1] C. Knell[1] A. Bosserhoff[2] E. Naschberger[3] M. Stürzl[3] M. Rübner[4] H. Seuss[5] M. Ruh[6] H.-U. Prokosch[1] B. Sedlmayr[1]

[1] Department of Medical Informatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[2] Institute of Biochemistry (Emil-Fischer-Center), Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[3] Division of Molecular and Experimental Surgery, Department of Surgery, Translational Research Center Erlangen, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[4] Department of Gynecology and Obstetrics, Comprehensive Cancer Center Erlangen-EMN, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[5] Department of Radiology, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[6] Department of Experimental Medicine 1, Nikolaus-Fiebiger-Center for Molecular Medicine, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Address for correspondence J. Christoph, MSc, Department of Medical Informatics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Wetterkreuz 13, Erlangen 91056, Germany (e-mail: jan.christoph@fau.de).

Appl Clin Inform 2017;8:1173–1183.

Abstract

Background Platforms like tranSMART assist researchers in analyzing clinical and corresponding omics data. Usability is an important, yet often overlooked, factor affecting the adoption and meaningful use. Analyses on the specific needs of translational researchers and considerations about the application of such platforms for education are rare.

Objectives The aim of this study was to test whether tranSMART can be used in education and how well medical students and professional researchers can handle it; to identify which kind of translational researchers—in terms of skills, experienced limitations, and available data—can take advantage of tranSMART; and to evaluate the usability and to generate recommendations for improvements.

Methods An online-based test has been done by medical students ($N = 109$) and researchers ($N = 26$). The test comprised 13 tasks in the context of four typical research scenarios based on experimental and clinical data. A web questionnaire was provided to identify both the needs and the conditions of research as well as to evaluate the system's usability based on the "*System Usability Scale*" (SUS).

Results Students and researchers were able to handle tranSMART well and coped with most scenarios: *cohort identification*, *data exploration*, *hypothesis generation*, and *hypothesis validation* were answered with a rate of correctness between 82 and 100%. Of the total, 72.2% of the teaching researchers considered tranSMART suitable for their lessons and 84.6% of the researchers considered the platform useful for their

Keywords

- translational research
- biomedical research
- surveys and questionnaires
- education
- adoption
- medical education
- computer systems evaluation

received
May 26, 2017
accepted after revision
October 23, 2017

DOI https://doi.org/10.4338/ACI-2017-05-RA-0085.
ISSN 1869-0327.

Copyright © 2017 Schattauer

License terms

daily work; 65.4% of the researchers named the nonavailability of a platform like tranSMART as a restriction on their research. The usability was rated "*acceptable*" with a SUS of 70.8.

Conclusion tranSMART is potentially suitable for education purposes and fits most of the needs of translational researchers. Improvements are needed on the presentation of analysis results and on the guidance of users through the analysis, especially to ensure the compliance of the analysis with the requirements of statistical testing.

Background and Significance

There is a rapid growing availability of biomolecular data such as genome, transcriptome, or proteome data (~*omics*[1]) which, for example, are obtained from samples of tumor tissue banks. This increases the need for translational researchers to integrate both the view and the analysis of the aforementioned data into clinical data, for example, to investigate the meaning of mutations or gene expressions for a clinical outcome.[2,3] Platforms like cBioPortal,[4] iDASH,[5] or tranSMART[6] provide a predefined set of methods for analysis and visualization such as *t*-tests, survival analyses, or heatmaps.[7] Although some of the platforms comprise an additional application programming interface (API) for programmers and statisticians, they are intended to be accessed via a browser by researchers with limited IT experience. Thus, *usability* is an important factor affecting both the adoption and the meaningful use of such tools.

Motivation and Related Work

While usability evaluations of health information technology receive much attention, there are only a few studies which address the usability of research platforms.[8–10] To our knowledge, there is no usability study for translational research platforms like iDASH, cBioPortal, or tranSMART. Some publications compare or explain such platforms only from the technical point of view.[7,11,12] Others describe their usage in studies[2,3,13–15] or in connection with other platforms or technologies.[16–18] This indicates a substantial need for further research.

Besides the availability of user-friendly software, "the promise of data in medicine and biology will not be realized without a new generation of students, researchers, and developers who are trained with state-of-the-art tools."[19] Knowledge and experience in several disciplines like medicine, bioinformatics, and molecular biology are often mentioned to be key factors for successful research in personalized medicine[20] and medical schools are requested to improve their curricula to include practical training for their applications.[21] However, the suitability of translational research platforms for education and medical student teaching is currently unknown. To close this gap in current research, we present an evaluation study of tranSMART as a representative example of translational research platforms.

The tranSMART Platform

The tranSMART platform enables an efficient exploration of data, presenting clinical and omics data integrated and easily accessible for further predefined analyses, data exploration, cohort identification, and the generation and validation of hypotheses. It has its roots in the i2b2 phenotype framework and consists of an entity-attribute-value store and a web frontend for interactive data exploration and analysis.[22] An active open-source community, organized by the tranSMART Foundation (recently joined to the *i2b2 tranSMART Foundation*), has further developed the program since 2013.[23]

Objectives

The objectives of our research were as follows:

- To test whether tranSMART is suitable for practical exercises of medical students and to examine how well future physicians can handle it.
- To identify which kind of translational researchers—in terms of skills, experienced limitations, and available data—can take advantage of tranSMART.
- To evaluate the usability of tranSMART in general.

Methods

The evaluation of tranSMART was conducted in Germany from November 2016 to January 2017 and comprised two steps: Step 1, an online exercise with corresponding online questions (*questionnaire part I*) and a subsequent option for feedback (free text) was given to 155 medical students to test whether tranSMART was suitable as an exercise for lectures of medical informatics. Moreover, it served as an upstream pretest for the next step. Step 2, nearly the same online exercise (questionnaire part I) but with an additional set of questions (*questionnaire part II*) was performed by 26 biomedical researchers to assess the researchers' correctness in handling tranSMART, their attitudes toward the system's usability, and their potential of using tranSMART for their own projects (see ►Fig. 1).

For the exercises of step 1 and step 2, we used tranSMART in its latest stable version (16.1) in a virtual machine (8 GB RAM and 4 CPUs) with Ubuntu 14.04 and a PostgreSQL 9.3 database. Based on the open-access data of the colorectal adenocarcinoma study COAD,[24] we created a dataset of 276 patients (28 clinical items and expression results from over 23,000 genes were assigned to each patient) with additional educational items to obtain the desired correlations and confounding variables. This dataset was imported into our instance of tranSMART using the tMDataLoader tool.[25] The test dataset contained all necessary data to solve the given tasks.

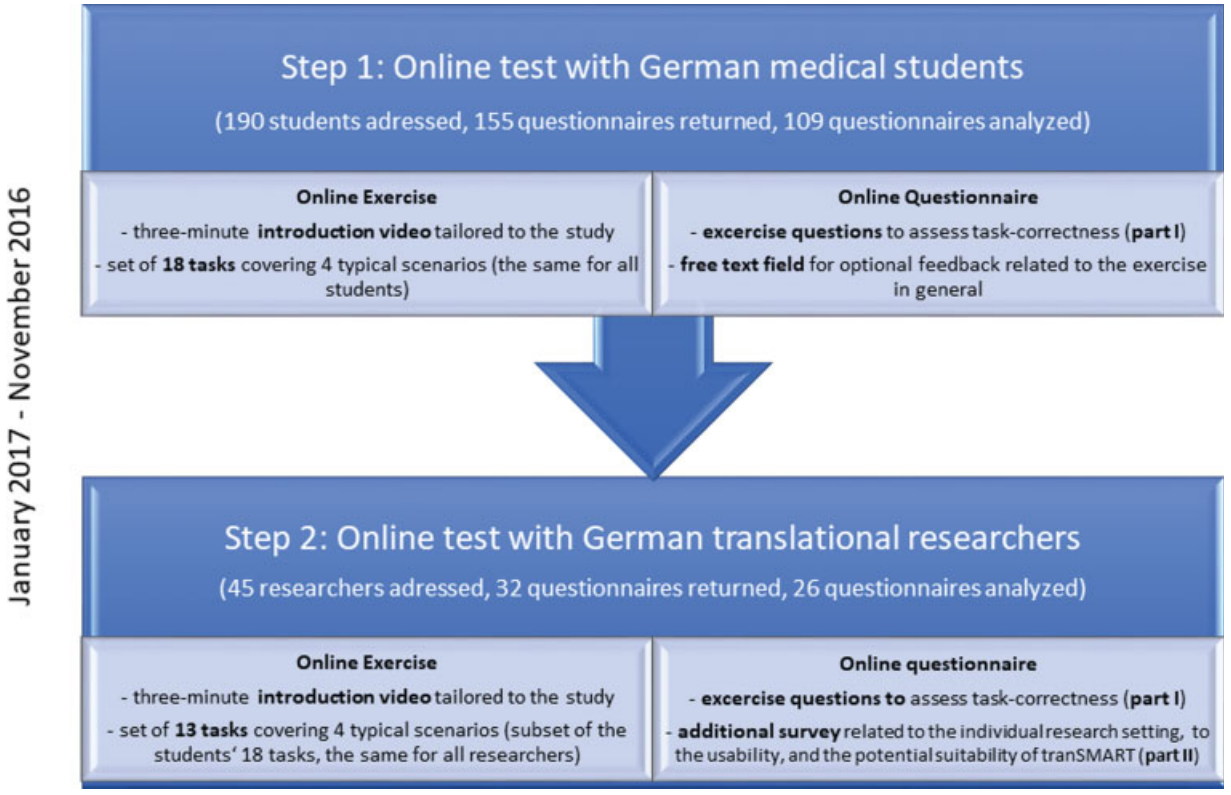

Fig. 1 Illustration of study.

Online Exercise with Exercise Questions (Questionnaire Part I: Students and Researchers)

A team comprised a computational molecular biologist, a computer scientist, a medical doctor with statistical expertise, and a usability expert developed an exercise with 18 tasks. Each task had to be solved by applying tranSMART and was represented in the form of a multiple-choice question. The tasks covered the four most important scenarios of tranSMART (according to literature[3]) and the most frequently used analysis methods in this context (according to a research group of our university medical center which has already been using tranSMART for more than 2 years[26]), see ►**Table 1**. All but one task could be answered correctly by a straight-forward application of tranSMART. However, one *statistics-aware question* required knowing that statistical tests can have limitations or preconditions and that it is not necessarily sufficient to have a significant *p*-value to accept a hypothesis. In this case, the hypothesis to check was "*smoking causes cirrhosis of the liver.*" The correlation is misleading because smoking and alcohol abuse is correlated and the latter is mostly responsible for the cirrhosis of the liver. Moreover, the provided analysis method (Fisher's test) of tranSMART could not check for causality. The students were requested to perform all 18 tasks. The researchers received only a subset of 13 tasks which was the same for all researchers (questionnaire part I, see ►**Supplementary Material 1**, available in the online version). Based on previous feedback and criticism of students (e.g., tasks were redundant or difficult to understand) 5 of the 18 tasks were removed from the entire

Table 1 Summary of the 13 exercises presented to the researchers (provided in detail in ►**Supplementary Material 1**, available in the online version)

Scenario	No. (tasks)	Example
Cohort identification	4×	How many men can be identified with a tumor in their rectum, but no metastases yet?
Data exploration	4×	Compare patients under the age of 70 with and without metastases —e.g., in terms of their affected lymph nodes.
Hypothesis generation	1×	Which gene expressions would you investigate further assuming an effect on the patient's survival time?
Hypothesis validation	3 + 1×	Test the hypothesis that the state "hypermutated" correlates with sex, using the Fisher's test.
Statistic-aware task		Test the hypothesis that smoking causes cirrhosis of the liver, using the Fisher's test.

task set and also excluded from any analysis. Researchers were offered the option of skipping individual tasks due to ambiguity of the software, technical problems, or other reasons (see ►Supplementary Material 1, available in the online version). We offered this option to reduce noise in the data and not to force the selection of an inappropriate answer.

Additional Questions about Usability and Suitability (Questionnaire Part II: Researchers)

In part II of the questionnaire in step 2 (for researchers only, see ►Supplementary Material 2), available in the online version), participants were asked to give basic demographic data as well as information about their actual research situation (31 questions) and to answer questions on the system's general usability, the suitability of tranSMART for the tasks, and the presentation of the information (72 questions). The questions regarding the system's general usability were based, among others, on the *System Usability Scale (SUS)*.[27] Closed questions were rated on a five-point Likert scale from *1 (do not agree at all)* to *5 (fully agree)*. Self-developed items were pretested for content and clarity by three researchers.

Procedure of the Study

As a first step, the online exercise with its 18 corresponding questions (part I of the questionnaire) was offered to all 190 medical students of a lecture on medical informatics (self-selected sample). The lecture was required for all medical students and planned to be held during the fifth semester (it is the first and only lecture in the curriculum related to computer science. Biometry is scheduled for the sixth semester). Participation in the exercise was optional but was motivated by a bonus that would be given for the subsequent written exam if at least 60% of the questions were answered correctly. At the beginning of the online exercise, the students were asked to give their informed consent for the scientific analysis of their results. The students were then instructed to watch a 3-minute introduction video (cf. http://youtube.com/watch?v=t4RVG7IYGF4) tailored to our study and the associated tasks about the handling of tranSMART. However, as we wanted to assess the usability and the ease of use of tranSMART, we did not give any additional training in advance. Next, the participants were requested to perform the 18 tasks related to the four typical scenarios of translational research and to answer the corresponding multiple-choice questions. Finally, they could give a general feedback on the exercise by a free text field. The time and the activity on the tranSMART server were logged for each student by assigning pseudonymized individual accounts.

As a second step, invitations for the participation in the online exercise and the survey were sent to 45 selected biomedical researchers from six different hospitals by email (self-selected sampling). The participation was voluntary and not financially compensated; anonymity was guaranteed. Like the students, the participating researchers were first instructed to watch the same introduction video and then told to perform 13 tasks and to answer the related multiple-choice questions (part I of the questionnaire). Additionally, they were asked questions on demographics, usability, and suitability (part II of the questionnaire).

All responses of the participants were automatically recorded and tabulated by the online service SoSci Survey.[28]

Data Analysis

The questionnaire was mainly analyzed descriptively. The total numbers and the percentages of each category were given for all rankings (categorical data). The means and the standard deviations (SDs) were calculated for numeric data (mainly demographic description). The total numbers and the percentage of correct answers were indicated for multiple-choice questions. Only those 13 tasks, which were given to both groups (students and researchers), were used for the analysis of the correctness of the tasks to render the groups principally comparable. For the items of the SUS, the score was calculated using Brooke's standard scoring method.[27] The textual feedback, which the students gave at the end of the exercise, was thematically categorized and rated as a "negative/neutral/positive statement" by two scientific assistants. If there was a disagreement, a consensus was achieved by discussion.

The correlation coefficients were determined for relations between personal variables and given assessments. All statistics were performed at a significance level of 5%. All calculations were performed using SPSS 23.0.

To exclude the results of students who might have only copied the exercise results from a fellow student, we excluded the questionnaires of those students from the analysis whose achieved correctness (according to the server logs) was incompatible with their activity or with their time on the tranSMART platform (e.g., no login into tranSMART at all or an online activity time in tranSMART of less than 30 minutes but more than 60% correctness).

Results

First, a *sample description* of the students and the researchers is given, followed by the results of the tranSMART exercise (part I of the questionnaire) which show *task correctness*. Then we present the results of part II which deals with the *suitability of tranSMART for different user groups* as well as the *usability*.

Sample Description

A total of 190 medical students were given the possibility to voluntarily participate in the tranSMART exercise. And 155 of these students performed this exercise and 133 of them, in turn, gave their consent to use their results for this publication. Twenty-four out of these 133 were excluded as they were suspected to have copied results. The remaining 109 students (57% response rate) had passed all their preliminary medical examinations; 70% of them were in the fifth semester, 25% were in their sixth semester, and 5% where in higher semesters. The sex ratio is female-biased (60:40%), which, however, is a typical gender relation for medical students at our university.

Thirty-two out of 45 addressed researchers took part in the online questionnaire and 26 of them completed the questionnaire and were therefore included in the analysis (58% of response rate). The researcher group comprised

participants of nine different specialties from four different German hospitals (see ►Supplementary Material 3 for more details, available in the online version).

The researchers had degrees in medicine (42.3%) or biology (30.8%); the main focus of their research areas was either molecular medicine/biology (17/26 = 65.4%) or medicine (7/26 = 26.9%). On average, the researchers had 8 years of work experience in their current specialty (SD: 7.9 years) and 9 years of research experience (SD: 8.6 years) including research time in previous specialties, if any. Three out of four respondents spent at least 60% of their working time doing research (20/26 = 76.9%) and two out of three researchers spent at least 10% of their time teaching (18/26 = 69.2%). At least one-third of the participants were involved in medical care in addition to their research projects. Most researchers rated their statistical knowledge (84.6%) as well as their knowledge of statistical software (65.4%) as *basic* and their computer experience as *average* (73.1%). A clear majority of respondents had no knowledge of programming or query languages like R or SQL (88.5%). Three participants (11.6%) had already used tranSMART in the past but rated their experience only as *basic knowledge*. None of the participants had previous experience with systems similar to tranSMART.

Task Correctness

►Table 2 shows to which extent students and researchers were able to solve the 13 tasks of the exercise. The results are subdivided into the four scenarios whereby the *statistic-aware question* has been considered as a separate aspect (which was not considered for calculation of the *task-correctness* of the scenario *hypothesis validation* to avoid an arbitrary bias). Both students and researchers showed similar results and gradations, although the students performed slightly better except in the scenario of *hypothesis generation*.

Most participants of both groups were able to solve the tasks of the scenarios *cohort identification*, *data exploration*, and *hypothesis generation* correctly (90–100%) by using the *summary statistics* function of tranSMART which was made for basic statistics. *Hypotheses validation*, the application and the interpretation of survival analyses and Fisher's tests, was solved correctly by 88% (students) and 82% (researchers), respectively. The researchers skipped 11.2% of all tasks: indicating ambiguity about the software for 4.7% of the tasks, technical problems for 0.6%, and other reasons for the remaining 5.9%.

In the *statistic-aware task*, the nonapplicability of the Fisher's test has not been detected by the majority of both students and researchers. Within the group of researchers, no significant (Fisher's exact test, $p = 1.000$) difference was observed between physicians (50% incorrect) and nonphysicians (55% incorrect). However, a moderate and significant correlation between a correct answer and the years of work experience (correlation coefficient = 0.48, $p = 0.028$) could be observed (see ►Table 3).

Suitability of tranSMART

Suitability of tranSMART for Students and Medical Education

►Table 4 shows the categorized feedback which was given by 90 students after the tranSMART exercise.

Table 2 Percentage of correct answers

Scenario	Correctness in % ($n/N/N^*$)		Missing of researchers
	Students	Researchers	
Cohort identification	95% (415/436/436)	90% (89/99/104)	5/104 c: 5×
Data exploration	96% (418/436/436)	91% (83/91/104)	13/104 a: 4×, b: 2×, c: 7×
Hypothesis generation	90% (98/109/109)	100% (21/21/26)	5/26 a: 4×, c: 1×
Hypothesis validation(without the statistic-aware task)	88% (288/327/327)	82% (56/68/78)	10/78 a: 6×, c: 4×
Statistic-aware task	50% (54/109/109)	38% (8/21/26)	5/26 a: 2×, c: 3×
Total	90% 1,273/1,417/1,417	86% 257/300/338	11.2% (38/338) a: 4.7% (16/338) b: 0.6% (2/338) c: 5.9% (20/338)

Abbreviations: n, number of correct answers; N, total number of valid answers; N^*, total number of answerable questions.
Skipping options: a = Not dealt with because of ambiguity of the software, b = Not dealt with because of technical problems, c = Not dealt with for other reasons. Example: In the scenario *cohort identification*, 109 students answered all four questions (one per task) ($109^*4 = 436 = N = N^*$) and were *415* times right (n) which corresponds to a rate of correctness of *95%*. The researchers answered in this scenario only *99* (N) out of *104* (N^*) tasks whereby the remaining 104–99 = *5* tasks have been skipped for other reasons (c) than ambiguity of the software or technical problems.

Table 3 Percentage of participants who answered the *statistic-aware question* correctly or confirmed the wrong hypothesis

Group	Recognized nonapplicability of the Fisher test (correct)	Confirmed hypotheses (incorrect)
Students	50% (54/109)	21% (23/109)
Researchers (all)	38% (8/21)	52% (11/21)
Physicians	30% (3/10)	50% (5/10)
Nonphysicians	45% (5/11)	55% (6/11)
Work experience >4 y)	70% (7/10)	30% (3/10)
Work experience ≤ 4 y)	9% (1/11)	73% (8/11)

Note: The remaining percentage chose the other two incorrect multiple-choice options. First breakdown of researchers into physicians (studied human medicine) and nonphysicians (any other training qualification), second breakdown according to their work experience (in years).

The majority of the students who mentioned the aspect of meaningfulness stated that this kind of exercise made sense for their studies (70.6% [24/34]). Only 14.7% (5/34) indicated that they recognized no benefit for their education. Nearly all students (93.8% [15/16]), who mentioned the aspect of fun, declared to have had fun performing the tasks with tranSMART. The comments about the grade of difficulty and the training material show that, on average, the level of difficulty was appropriate (about one-third found it *a little bit too difficult*, one-third found the grade to be *ok*, and the last third felt *quite comfortable* with it). The training material, especially the introduction video, was considered helpful. However, the required knowledge of statistics was assessed as partially too high. According to the results of the questionnaire part II, 72.2% (13/17) of the researchers who were also involved in teaching at that time stated that tranSMART would be useful for their lessons, and a further 17.6% (3/17) agreed partially. Including the nine researchers who were not teaching at that time, results show that a similar rate of 73.1% (19/26) of the participants agreed (largely or fully) and further 19.2% (5/26) agreed partially—as shown in ►Fig. 2 and ►Supplementary Material 4, available in the online version.

Table 4 Categorized free-text feedback of 90 students

Category	Negative	Neutral	Positive
Exercise made sense for my studies	5	5	24
Exercise was fun	0	1	15
Grade of difficulty was fine	5	4	4
Prerequisite of statistic knowledge was appropriate	5	1	2
Training material was helpful	5	3	10
Tasks were properly explained	6	3	10
Working time was appropriate	5	1	2

Note: Thirty-four out of 90 students made a statement on the perceived meaningfulness of the exercise and 24 of them see some sense in the exercise.

Suitability of tranSMART for Biomedical Researchers and Their Demands

As shown in ►Table 5, nearly two-thirds of the researchers (17/26 = 65.4%) indicated in the survey that the current lack of appropriate research platforms in their research environment is one of their research constraints. Missing facilities for the analysis of molecular-biological raw data (61.5%) and difficult access to clinical patient data (53.8%) were named as further relevant barriers.

After each of the researchers had performed the 13 tasks of the tranSMART exercise within 45 to 60 minutes, they rated the platform. Accordingly, 84.6% (22/26) of all interviewed researchers agreed largely or fully with the statement that tranSMART could be useful for their work (see ►Fig. 2). About half of all researchers, 53.8% (14/26), would prefer tranSMART to their traditional way of analyzing data.

The time spent on different tasks was rated as adequate (almost all participants gave a positive answer for all tasks), whereby the *comparison of cohorts* got the best rating.

Within the subset of biomedical researchers who spent at least 60% of their working time on research, 95% (19/20) considered tranSMART to be fully (70%) or largely (25%) useful for their research. Within the small group of three researchers who had knowledge in programming or query languages such as R or SQL, two agreed largely and one agreed partially to the statement that tranSMART would be useful for their work.

According to a question about the need for analysis methods, the *correlation analysis* would be most frequently used (84.6% of the participants), followed by the *t-test/chi-square test* (80.8%). In addition, the majority of researchers would make use of the *Fisher's test*, *ANOVA* analysis, *survival analysis* (70% each), and *heatmaps* (61.5%). Between 30 and 50% of the researchers wanted to use *linear regression* (50.5%), *clustering* (38.5%), and *array-based comparative genomic hybridization (aCGH) analysis* (38.5%). *Principle component analysis* and *GWAS* (genome-wide association studies;[29] 20% each) would be the methods of analysis which would be used least of all.

GSEA (*gene set enrichment analysis*;[30] 34.5%) as well as *functional pathway analysis* (one free text remark) were mentioned as needed but are not yet available in tranSMART.

When they were asked for the kind of data they would need on a research platform like tranSMART (see ►Supplementary Material 5, available in the online version), most participants (84.6%) reported that they would need patient data (medical history, diagnoses, therapies) for their research. Approximately 70% of the respondents would require *mRNA* data, *miRNA* data, *DNA methylation* data, *image data*, and *proteomic* data. About half of them would need *metabolomic* data (57%), *SNP/SNV* (50%), or

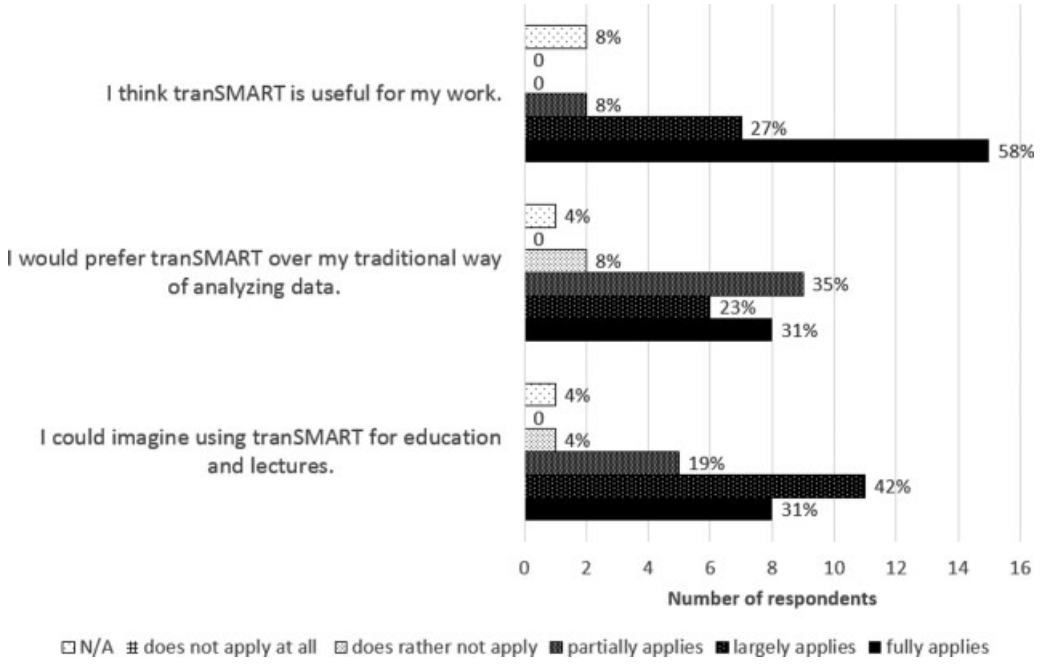

Fig. 2 Assessment of the suitability of tranSMART for research and education. The bars represent the absolute numbers of the respondents shown on the *x*-axis; the bar labels show the percentage of the respondents (100% corresponds to all 26 researchers). The category "does not apply at all" has never been selected.

CNV (42%). Except *DNA methylation* data, all required data types are currently supported by tranSMART.

Not only the two participating radiologists but also 62.5% (15/24) of the other researchers agreed at least partially to need image data for their research. Three-fourths (75%) of the researching participants with a focus on medicine rather than on molecular medicine/biology required omics data likewise (especially *mRNA*, *miRNA*, and *DNA methylation*).

Usability Assessment

The results of the "System Usability Questionnaire" show an average score of 70.8 (*ok*/*acceptable* usability)—ranging from 35 (*poor*) to 95 (*excellent*). ►Fig. 3 shows the distribution of the individual SUS scores classified by usability categories according to Bangor et al.[31] A further correlation analysis revealed a significant and positive correlation between the individual SUS scores and the years of work experience in research (Pearson: 0.931, $p < 0.001$).

When they were asked for their opinion on the quality of presented information, most respondents stated to be fully or largely satisfied with the interface design, the representation of colors, the modification of diagrams, the navigation, the way of representing results, and the font size. The least satisfaction was expressed for the item *terminology of options*. Further results showed that the participants were divided regarding their opinion about the error tolerance of the system (defined as possibility to achieve accurate results with no or only minimal corrective action, see ►Table 6).

Table 5 Level of agreement to the statement "*the following circumstances restrict my research*"

Kind of restriction	Not at all	Rather not	Partially	Largely	Fully	Missing	Acceptance	Rejection
Missing platforms for research such as tranSMART	2	2	3	8	6	5	17 **65.4%**	4 **15.4%**
Insufficient facilities to analyze biomedical raw data	4	1	1	9	6	5	16 **61.5%.**	5 **19.2%**
Insufficient access to medical data of patient care	2	3	4	7	3	7	14 **53.8%**	5 **19.2%**
Too little staff due to insufficient funding	1	4	4	6	2	9	12 **46.2%**	5 **19.2%**
Insufficient facilities to analyze biomedical samples	5	4	4	4	3	6	11 **42.3%**	9 **34.6%**
Too little staff due to the lack of qualified candidates	6	4	6	1	0	9	7 **26.9%**	10 **38.5%**

Note: Acceptance sums up the levels of (applies) *partially, largely,* and *fully. Rejection* sums up the levels of *does rather not apply* and *applies not at all*. The table shows the absolute number of the researchers who chose the respective option.

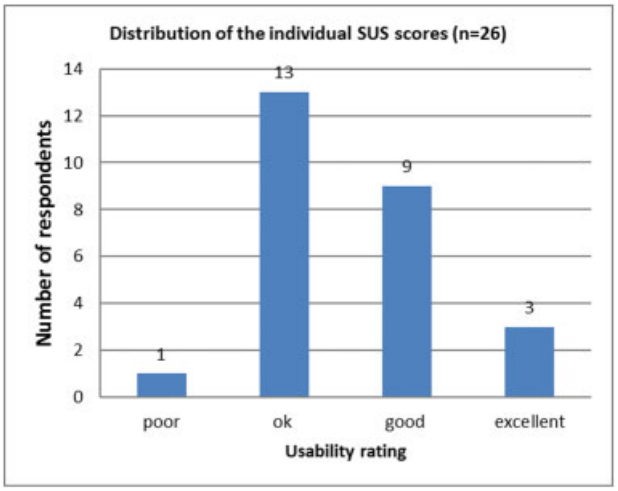

Fig. 3 Distribution of the individual System Usability Scale (SUS) scores ($N = 26$).

About 70% of the respondents preferred to have additional information about the application of statistical methods or tests; 57.7% of the researchers would prefer the display of additional information about specific options/parameters. Almost all participants require additional functionalities, in particular a link to the patient record (76% fully/rather agree), a link to external gene databases (92.3% fully/rather agree), and a link to external network/path databases (84.6% fully/rather agree).

In open questions, the physicians valued user friendliness of the system the most ($n = 10$) and a clear overview of statistical results and diagrams ($n = 6$). The individualization of the presentation of diagrams (colors, text fields, $n = 4$) and the addition of further information for tests and analyses ($n = 3$) still leave room for improvement.

Table 6 Rating of the researchers' satisfaction concerning different usability aspects (in percent)[a]

Satisfied with	Fully/ Largely	Partially	Rather not/Not at all
Interface design	61.6	26.9	11.5
Color representation of diagrams	57.7	26.9	15.4
Modification of diagrams	69.2	15.4	15.4
Navigation	73.1	19.2	7.7
Way of presenting results	57.7	38.5	3.8
Font size	80.8	11.5	7.7
Terminology of options	34.6	30.8	34.6
Error tolerance	19.2	50.0	26.9

[a]The remaining percentages in some cases are caused by missing answers.

Discussion

The aim of this study was (1) to investigate whether tranSMART is suitable for practical exercises and for the education of future physicians, (2) to identify which kind of translational researcher can take advantage of tranSMART and what the demands on this platform are like, and (3) to evaluate the usability of tranSMART.

Discussion of Results

Suitability of tranSMART for Students and Medical Education

The medical students were largely able to handle the four scenarios very well. Since the students dealt with all tasks for about 1 hour on average, the user experience was neither extensively profound nor superficial. The free text feedback indicated that the majority liked the exercise and assumed that it would make sense for their studies or even stated that it stimulated them for a thesis in medical statistics. Therefore, we conclude that tranSMART is well suited to be integrated into the curriculum of medical students as a practical exercise for the aspects of translational research. Ideally, the integration would be parallel or subsequent to biometrics lessons. This assumption is confirmed by the fact that 72.2% of the researchers who are involved in teaching agreed that they considered tranSMART to be useful for their lectures and exercises.

Suitability of tranSMART for Biomedical Researchers and Their Demands

Similar to the students, the researchers solved the scenarios of *cohort identification*, *data exploration*, and *hypothesis generation* with over 90% of correct tasks, whereas the scenario of *hypothesis validations* showed a correctness of only 82% and thus seems to have been more challenging.

Concerning the statistic-aware task, it was remarkable that even 50% of the researchers with a strong medical background confirmed a medical nonsense hypothesis. We assume that a lack of time or overlooking the key word "causing" may have been the cause. Although in a real-world application this error would perhaps be detected in subsequent reflections, it suggests that the requirements for statistical methods are not always sufficiently taken into account by the user and that there is a risk that the interpretation of the results might be performed too thoughtlessly. For these reasons, it would be highly desirable to implement all test conditions which can automatically be checked before the start of the analysis (e.g., sample size, Gaussian's distribution) and which, in case of doubt, would show warnings or refuse the test run. We additionally recommend the provision of context-sensitive information about the application and the background of statistical methods for the user if requested. This has also been stressed by Bauer et al.[14]

The majority of the researchers considered tranSMART to be useful for their daily work and stated that the missing availability of a research platform such as tranSMART remains a potential restriction for their research. Our results show that the need for such a platform is very high (at least in case of scenarios/tasks of this kind and with a patient cohort

of at least 10–20 patients)—independent of the participants' experience or focus. Researchers with a background in programming or query languages stated as well that they would benefit from tranSMART. This was an unexpected result for us because queries from a command line or analyses by scripts offer greater possibilities and more detailed results than queries which are run through the web interface of tranSMART. Yet the expected usefulness of a platform like tranSMART may be overestimated because its ultimate potential depends on the completeness and the quality of the raw data to be investigated.

The researchers indicated that the methods which are offered by tranSMART cover the daily need for tools even if some participants missed the function of the GSEA and the functional pathway analysis. This shows the need for the implementation of these methods according to the usability criteria. The implementation might be done in the form of a dynamic and interactive analysis workflow within the smartR-plugin[32] (parallel to our implementation of the survival analysis[33]).

However, as 20% of the researchers marked GWAS or GSEA as unknown methods, we conclude that they would additionally benefit from a future training in the applicability of these analysis methods.

The data types which were supported by tranSMART largely correspond to the needs of the researchers except for the *DNA-methylation* data type. The latter is considered to be one of the four most important data types, but it is not yet supported in tranSMART. We therefore recommend the provision of this functionality in future.

Usability

The overall usability (SUS score) of tranSMART was assessed as *acceptable*, which means that most researchers were satisfied with the ease of use and the learnability of tranSMART. However, this result also indicates that there is room for improvement. This is as well reflected by the answers to the open questions which demonstrate that the researchers, above all, need additional information about parameters or options and further information about statistical tests or methods for their work (e.g., test preconditions). Considering the average SUS score of 71, our results are in line with the findings of studies of other research platforms: Mathew et al,[10] for example, revealed a mean SUS score of 67 for a collaborative research platform and Wozney et al[8] reported an overall mean score of 70 for their "Intelligent Research and Intervention Software," which is used to deliver psychosocial interventions across a range of clinical settings. To our knowledge, there is no study which addresses the usability aspects of translational research platforms. So far it has always been claimed that translational research platforms are user friendly (e.g., "[tranSMART] can be used by users […] without any expertise in computer science."[3]), but proof has not yet been provided. Sinha and Markatou,[34] for example, designed a platform "with a focus on usability and interpretability of analysis for the researcher" but did not evaluate its usability. Cerami et al[35] described an open platform for exploring cancer genomics data and stated that the system is intuitive for researchers and that a "key feature of the…portal is ease of use." However, they presented no results that would support such statements. We successfully evaluated tranSMART with a positive result and thus contributed to close this research gap.

Limitations and Future Work

Our study has several limitations:

- We excluded students who were reasonably suspected to have copied. The exclusion criteria were designed to avoid an overestimation of the students' correctness but may have excluded a small number of participants who used tranSMART very well and quickly.
- We tested tranSMART only with a medium-sized sample of students and clinical researchers and most participants were from a single center, which limits the transferability of the results—especially in terms of suitability—to other settings (e.g., other institutions/hospitals).
- The study was preceded by a pre-test to assure a high quality of data, the appropriateness of the questions, and the technical functionality of the online questionnaire. However, online testing does not allow for the control of environmental conditions (e.g., distractions), which might have influenced the participants' results.
- Other software programs like cBioPortal or iDASH offer similar functions to those of tranSMART but were not tested in comparison. Results which are very specific to tranSMART, such as the assessed usability score SUS, cannot be transferred to them. Nevertheless, the stimulation to investigate and consider the usability as well as the needs of the researchers concerning the required data and the methods of analysis may also apply to them.
- To test tranSMART, we developed a comprehensive set of test tasks, which cover the main functionalities of tranSMART, but there are also some analysis features for programmers which are more advanced—such as the hierarchical clustering or the access of data through an API. The introduction of video for all participants only covered the use of data through the web frontend (data analysis/interpretation) and was tailored to our study. The official tranSMART Web site offers training for researchers which comprises multiple videos whereby each video lasts about an hour so that the complete training would take several hours.
- The evaluation of the platform comprised the task area of "data analysis/interpretation." Other aspects, especially data modeling and data import (e.g., from public databases like GEO[36] or TCGA[37]), were not tested. To cover the task areas of data modeling and data import, tranSMART and similar platforms do not yet address clinical researchers but mainly computer scientists who perform these tasks in consultation with the researchers.

Future work should include a component that evaluates which parts of the curriculum of medical students can be covered by the use of tranSMART. Furthermore, there should be a study which evaluates the efficacy of the transfer of competences in research methodology and biomedical research by the application of tranSMART (compared with other currently available options). These evaluations should also include the viewpoints of teachers in the fields of

medicine. In addition, it would be necessary to evaluate how to embed intelligent prompting the use of appropriate statistical methodology into the respective software, which is a persistent weakness among translational researchers.

Conclusion

By investigating the suitability and usability of tranSMART for education and translational research, we have addressed two aspects which have been rarely considered so far and we have extended the applications of tranSMART. We hope that our recommendations will lead to a further refinement and expansion of the platform's services to offer its users even better support for high-content analyses. Furthermore, the evaluation framework (the tasks and the online questionnaire) can be used by the developers of similar research platform as a blueprint or as an input for their own evaluation activities. Finally, we recommend continued user evaluation of the software which is used by health researchers to encourage future improvements.

Clinical Relevance Statement

The present work provides new insights into translational research platforms by taking the example of tranSMART and by considering its suitability for research and education and its usability in general. It gives explicit recommendations to both users and developers with the primary aim of enhancing usability and minimizing statistical misinterpretations. This might help ensure a wider and better applicability of such platforms with the result of a more efficient and more qualitative translational research.

Multiple Choice Questions

1. What kind of support does tranSMART offer to avoid statistical misinterpretations?

   A. Checks for the appropriate kind of distribution (e.g., normal distribution)
   B. Checks for a sufficient sample size
   C. Context-sensitive information about the test requirements
   D. None

   Correct Answer: The correct answer is D. As partly mentioned in the sections of "Results" and "Discussion," the lack of support to avoid statistical misinterpretations is one of the points, which has been mainly criticized. This is demonstrated by the example of the statistic-aware task.
2. Which two groups are in the primary focus of this publication?

   A. Researchers and developers
   B. Researchers and medical students
   C. Medical students and administrators
   D. Physicians and students of computer science

   Correct Answer: The correct answer is B. Although there are also some suggestions for developers, the clear focus of the presented work is on researchers and medical students.

Protection of Human and Animal Subjects

Our research plan was reviewed by the Ethics Committee of the Friedrich-Alexander-Universität Erlangen-Nürnberg which considered this type of research to be exempted from its approval process and issued a written confirmation thereof. In addition, we declare that the study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects.

Conflict of Interest

None.

Acknowledgments

The researchers who participated in this study are gratefully acknowledged for their precious time and effort. We also thank S. Newe and A. Purohit for linguistically proofreading the manuscript. The research has been supported by the German Federal Ministry of Education and Research (01ZZ1606H) as well as by the Smart Data Program of the German Federal Ministry for Economic Affairs and Energy (1MT14001B). The present work was performed in fulfillment of the requirements for obtaining the degree "Dr. rer. biol. hum." from the Friedrich-Alexander-Universität Erlangen-Nürnberg (J.C.).

References

1 Joyce AR, Palsson BØ. The model organism as a system: integrating 'omics' data sets. Nat Rev Mol Cell Biol 2006;7(03):198–210
2 Jonnagaddala J, Croucher JL, Jue TR, et al. Integration and analysis of heterogeneous colorectal cancer data for translational research. Stud Health Technol Inform 2016;225:387–391
3 Rance B, Canuel V, Countouris H, Laurent-Puig P, Burgun A. Integrating heterogeneous biomedical data for cancer research: the CARPEM infrastructure. Appl Clin Inform 2016;7(02):260–274
4 Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;6(269):pl1
5 Ohno-Machado L, Bafna V, Boxwala AA, et al; iDASH team. iDASH: integrating data for analysis, anonymization, and sharing. J Am Med Inform Assoc 2012;19(02):196–201
6 Athey BD, Braxenthaler M, Haas M, Guo Y. tranSMART: an open source and community-driven informatics and data sharing platform for clinical and translational research. AMIA Jt Summits Transl Sci Proc 2013;2013:6–8
7 Canuel V, Rance B, Avillach P, Degoulet P, Burgun A. Translational research platforms integrating clinical and omics data: a review of publicly available solutions. Brief Bioinform 2015;16(02): 280–290
8 Wozney L, McGrath PJ, Newton A, et al. Usability, learnability and performance evaluation of Intelligent Research and Intervention Software: a delivery platform for eHealth interventions. Health Informatics J 2016;22(03):730–743
9 Soto-Rey I, N'Dja A, Cunningham J, et al. User satisfaction evaluation of the EHR4CR Query Builder: a multisite patient count cohort system. BioMed Res Int 2015;2015:801436
10 Mathew D, McKibbon KA, Lokker C, Colquhoun H. Engaging with a Wiki related to knowledge translation: a survey of WhatisKT Wiki users. J Med Internet Res 2014;16(01):e21
11 Dunn W Jr, Burgun A, Krebs M-O, Rance B. Exploring and visualizing multidimensional data in translational research platforms. Brief Bioinform 2016:bbw080

12 Wheelock CE, Goss VM, Balgoma D, et al; U-BIOPRED Study Group. Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. Eur Respir J 2013;42(03): 802–825
13 Bierkens M, van der Linden W, Weistra W, et al. Abstract 3166: querying, viewing and analyzing colorectal cancer translational research studies in tranSMART. Cancer Res 2016;76(14, Suppl):3166
14 Bauer CR, Knecht C, Fretter C, et al. Interdisciplinary approach towards a systems medicine toolbox using the example of inflammatory diseases. Brief Bioinform 2017;18(03):479–487
15 Visser PJ, Streffer J. A European medical information framework for Alzheimer's disease (EMIF-AD). Alzheimers Dement 2015;11 (07):120–121
16 Camacho Rodriguez JC, Stäubert S, Löbe M. Automated import of clinical data from HL7 messages into OpenClinica and tranSMART using Mirth Connect. Stud Health Technol Inform 2016; 228:317–321
17 Firnkorn D, Merker S, Ganzinger M, Muley T, Knaup P. Unlocking data for statistical analyses and data mining: generic case extraction of clinical items from i2b2 and tranSMART. Stud Health Technol Inform 2016;228:567–571
18 Wang S, Pandis I, Wu C, et al. High dimensional biological data retrieval optimization with NoSQL technology. BMC Genomics 2014;15(Suppl 8):S3
19 Huang K, Liu Y, Huang Y, et al. Intelligent biology and medicine in 2015: advancing interdisciplinary education, collaboration, and data science. BMC Genomics 2016;17(Suppl 7):524
20 Gomez-Cabrero D, Abugessaisa I, Maier D, et al. Data integration in the era of omics: current and future challenges. BMC Syst Biol 2014;8(Suppl 2):I1
21 Salari K. The dawning era of personalized medicine exposes a gap in medical education. PLoS Med 2009;6(08):e1000138
22 Kohane IS, Churchill SE, Murphy SN. A translational engine at the national scale: informatics for integrating biology and the bedside. J Am Med Inform Assoc 2012;19(02):181–185
23 i2b2 tranSMART Foundation. The i2b2 tranSMART Foundation announce completion of Merger [cited 2017 Aug 18]. Available from: http://transmartfoundation.org/the-i2b2-transmart-foundation-announce-completion-of-merger/
24 Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487 (7407):330–337
25 Alexander Bondarev. tMDataLoader [cited 2017 Mar 15]. Available from: https://github.com/ThomsonReuters-LSPS/tMDataLoader
26 Christoph J, Knell C, Naschberger E, et al. Two years of tranSMART in a university hospital for translational research and education. Stud Health Technol Inform 2017;236:70–79
27 Brooke J. SUS-A quick and dirty usability scale. Usability Eval Industry 1996;189(194):4–7
28 Leiner DJ. SoSci Survey-Online Questionnaire [cited 2017 Mar 15]. Available at: https://www.soscisurvey.de/
29 Bush WS, Moore JH. Chapter 11: genome-wide association studies. PLOS Comput Biol 2012;8(12):e1002822
30 Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102(43): 15545–15550
31 Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008;24(06): 574–594
32 Herzinger S, Gu W, Satagopam V, et al; eTRIKS Consortium. SmartR: an open-source platform for interactive visual analytics for translational research data. Bioinformatics 2017;33(14): 2229–2231
33 Knell C, Sedlmayr M, Christoph J. Developing interactive plug-ins for tranSMART using the SmartR framework: the case of survival analysis. Stud Health Technol Inform 2017;236:375–382
34 Sinha A, Markatou M. A platform for processing expression of short time series (PESTS). BMC Bioinformatics 2011;12:13
35 Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2(05):401–404
36 Barrett T, Troup DB, Wilhite SE, et al. NCBI GEO: archive for high-throughput functional genomic data. Nucleic Acids Res 2009;37 (Database issue):D885–D890
37 Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015;19(1A):A68–A77