Session II: Praktische Arbeit mit tranSMART

Aufbereitung und Analyse von Daten in phänotypischen und genotypischen Forschungsdatenbanken mit tranSMART | 05.08.2016 | Berlin

Teil 1

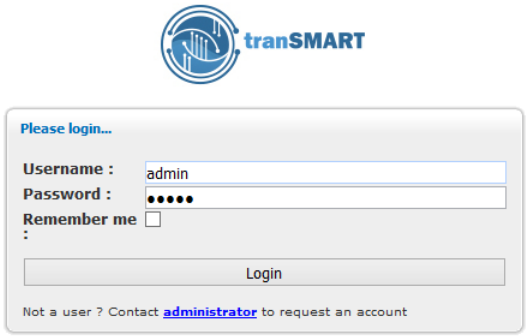
Ein kurzes einführendes Szenario mit tranSMART

Aufgabe 1.1

Als motivierendes Beispiel soll ein Signifikanztest (exakter Fisher-Yates-Test oder exakter Chi-Quadrat-Test) an einem Datensatz aus einer Brustkrebsstudie durchgeführt werden. Ein Fishertest liefert ein Maß für die Unabhängigkeit zweier Merkmale.

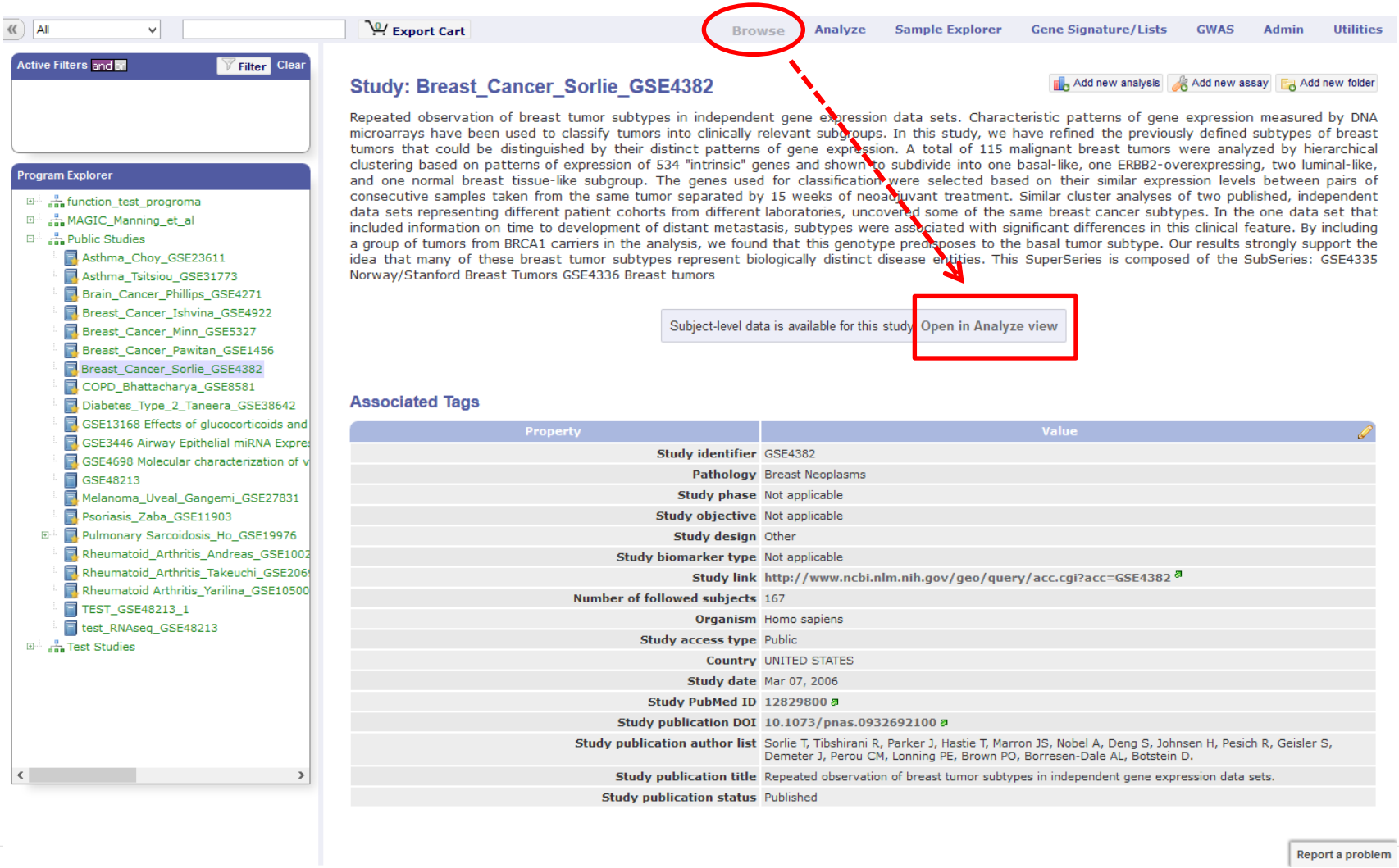
- Wie unabhängig ist die Überlebenszeit von der Tumorgröße?

Anmelden am System

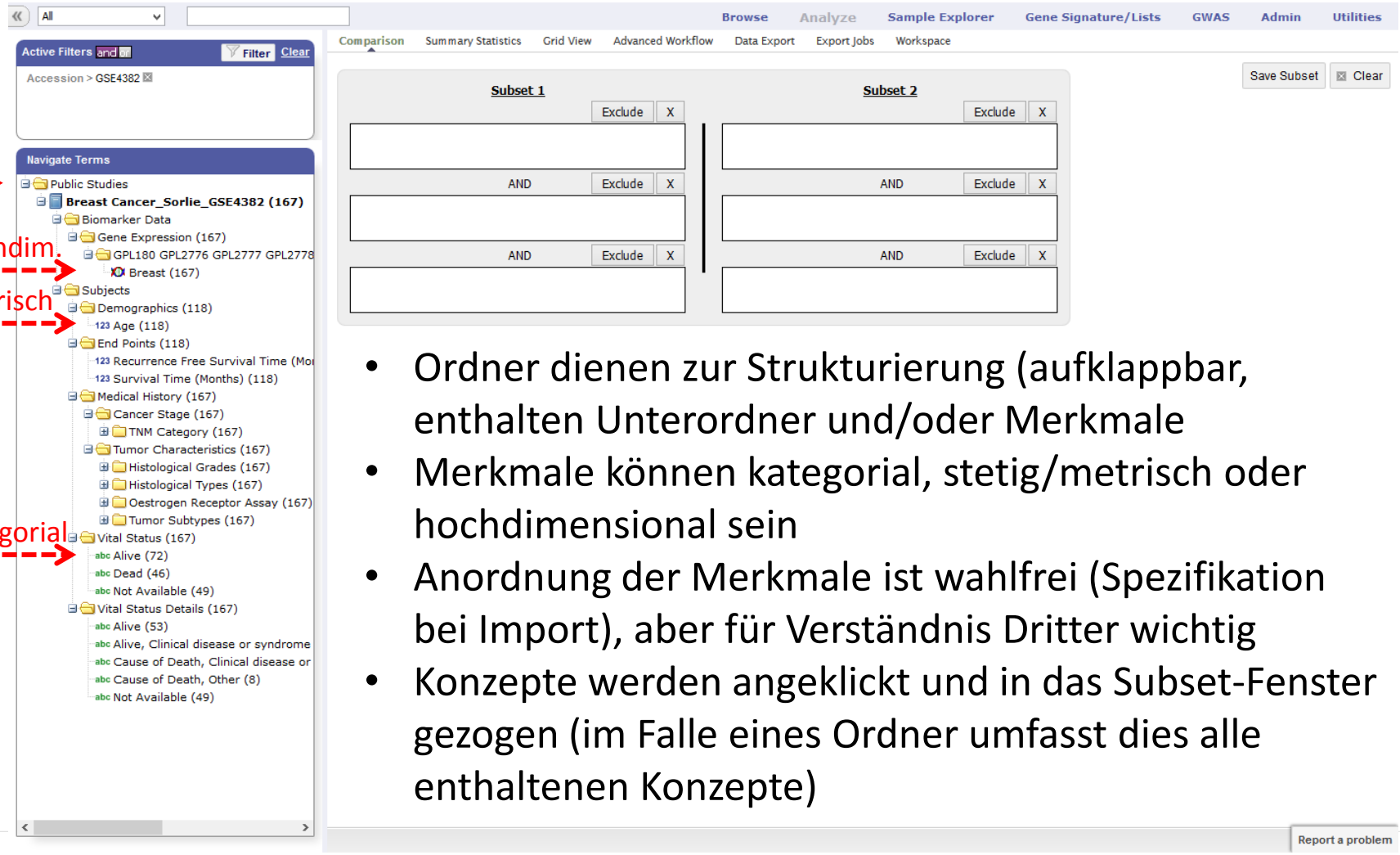


- Instanz 1:
  - URL: http://postgres-demo.transmartfoundation.org/transmart
  - uid: admin
  - pwd: admin
- Instanz 2:
  - URL: http://public.etriks.org
  - uid: guest
  - pwd: transmart2015

Browse: Studienübersicht und Metadaten

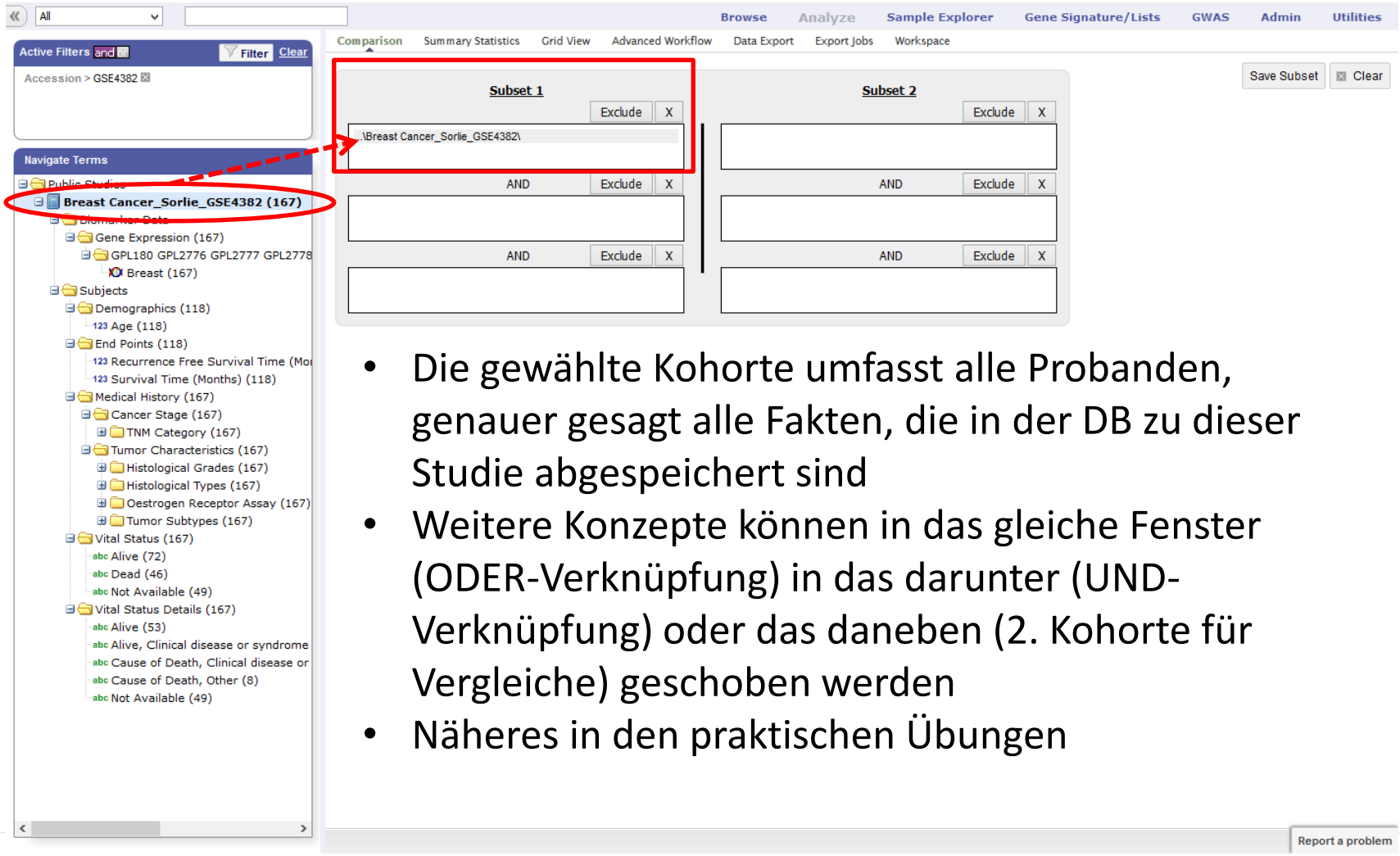


Links: Navigationsbaum für klinische Konzepte



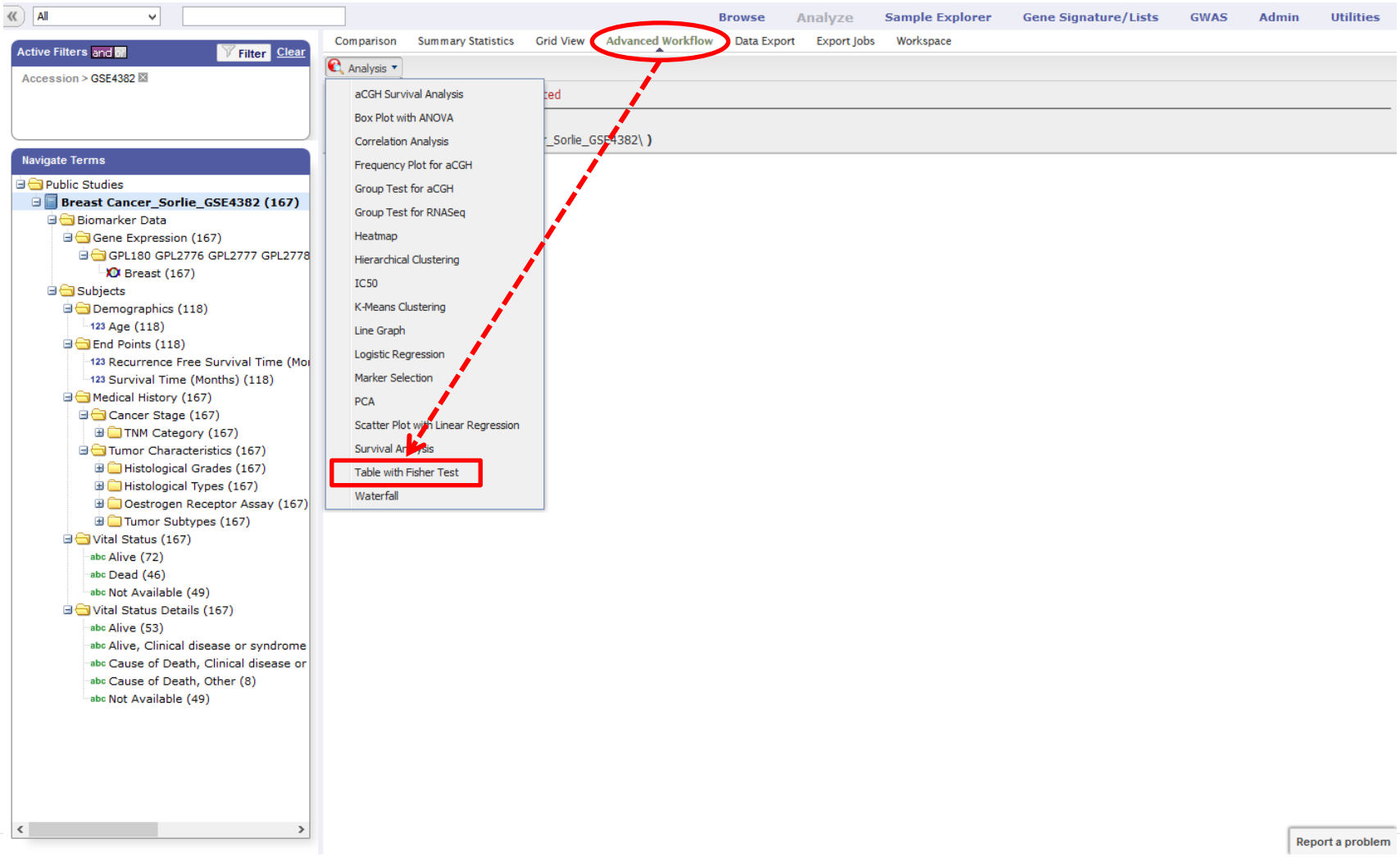
- Ordner dienen zur Strukturierung (aufklappbar, enthalten Unterordner und/oder Merkmale
- Merkmale können kategorial, stetig/metrisch oder hochdimensional sein
- Anordnung der Merkmale ist wahlfrei (Spezifikation bei Import), aber für Verständnis Dritter wichtig
- Konzepte werden angeklickt und in das Subset-Fenster gezogen (im Falle eines Ordner umfasst dies alle enthaltenen Konzepte)

Auswahl der gesamten Studie

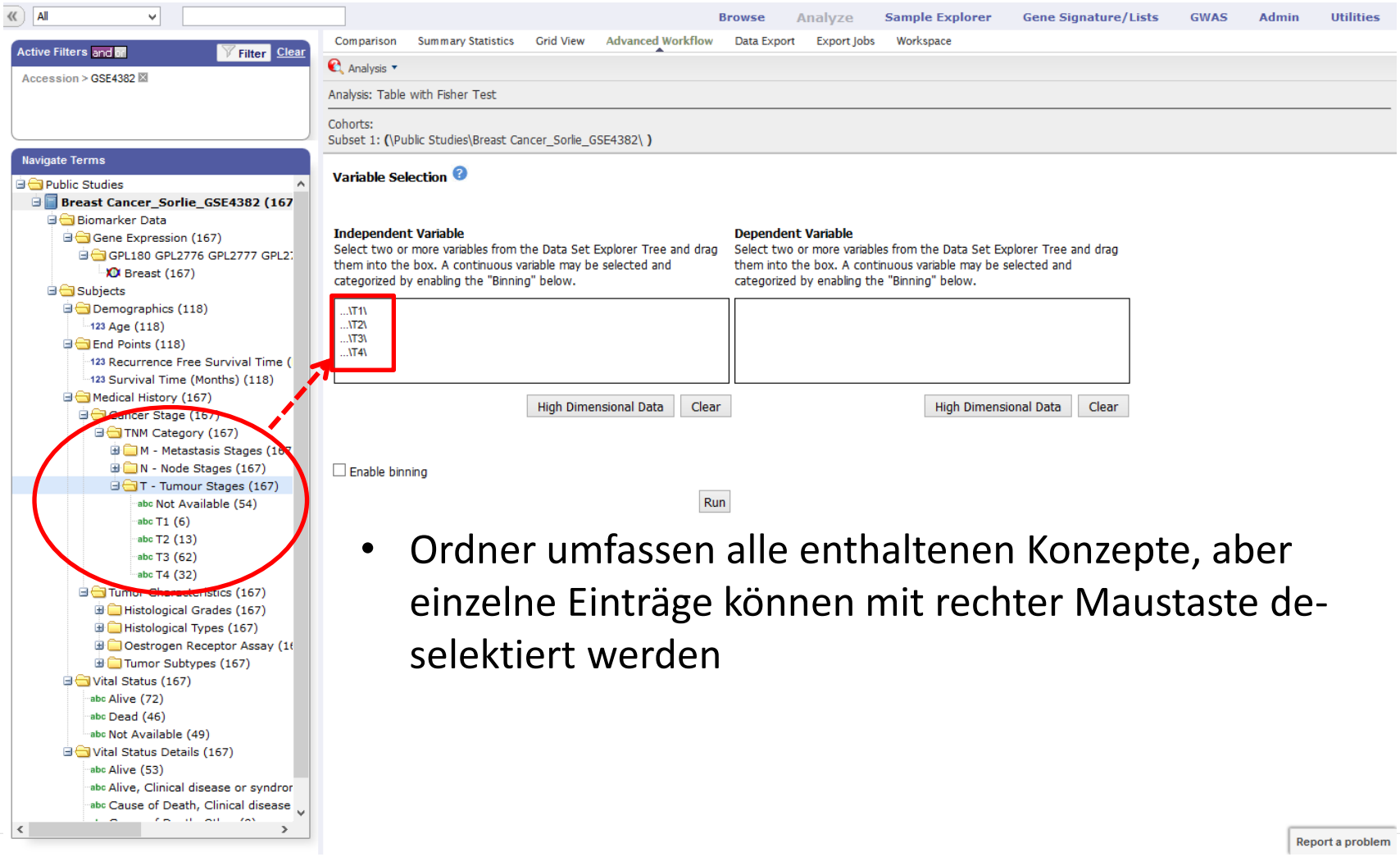


- Die gewählte Kohorte umfasst alle Probanden, genauer gesagt alle Fakten, die in der DB zu dieser Studie abgespeichert sind
- Weitere Konzepte können in das gleiche Fenster (ODER-Verknüpfung) in das darunter (UND-Verknüpfung) oder das daneben (2. Kohorte für Vergleiche) geschoben werden
- Näheres in den praktischen Übungen

Gewünschte Analysefunktion auswählen

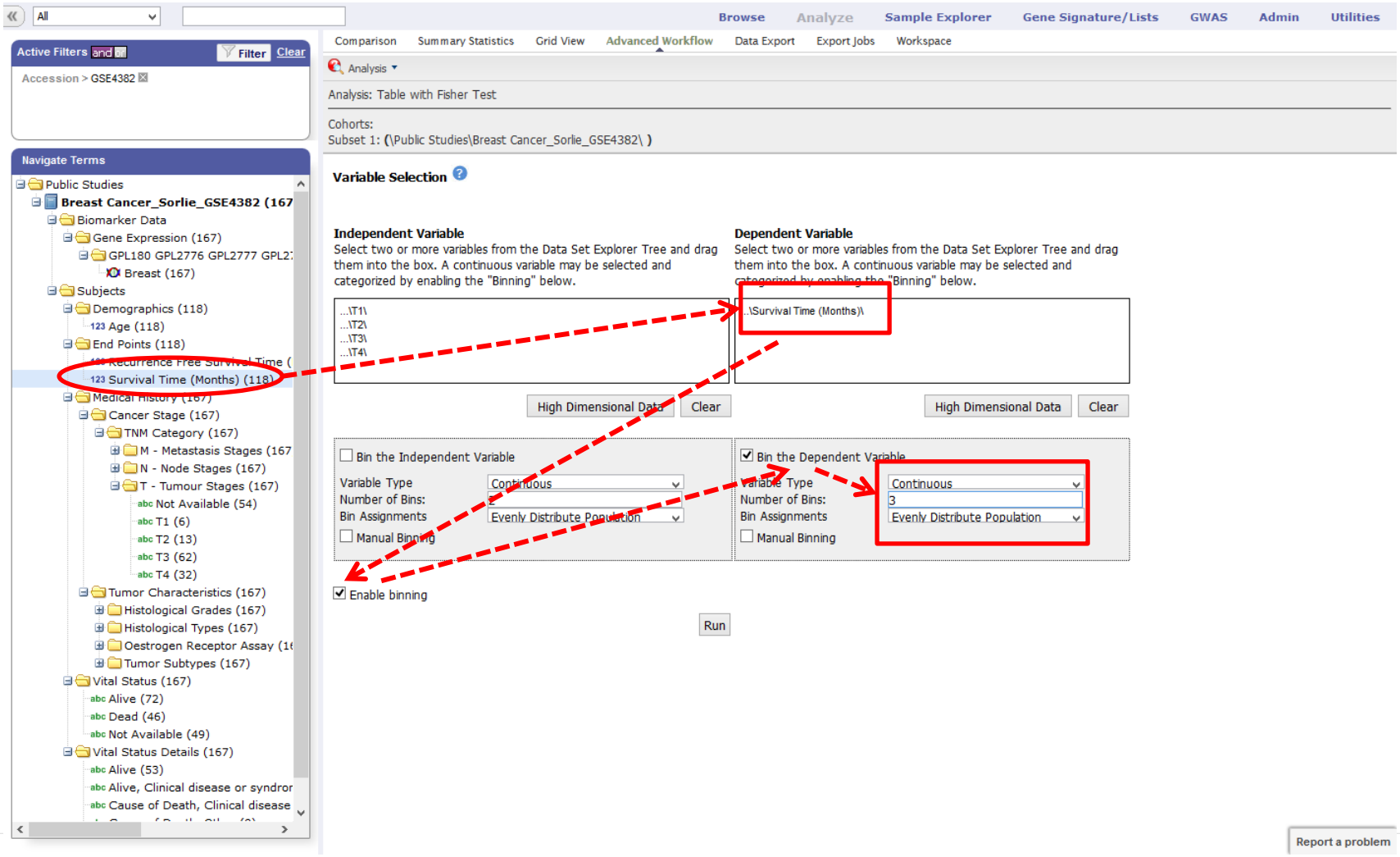


Es müssen die zwei zu untersuchenden Konzepte ausgewählt werden: 1. Tumorgröße nach TNM

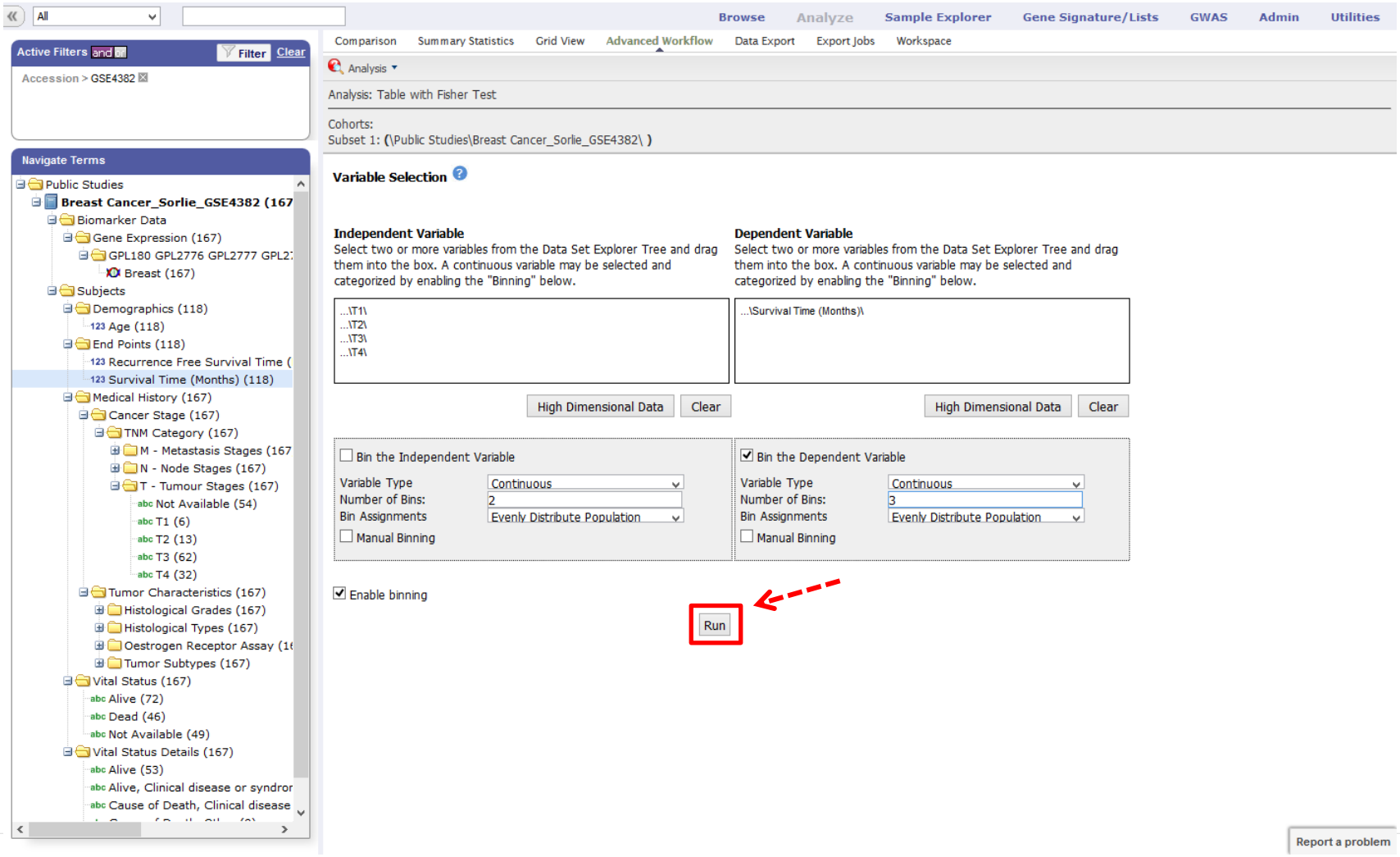


- Ordner umfassen alle enthaltenen Konzepte, aber einzelne Einträge können mit rechter Maustaste de-selektiert werden

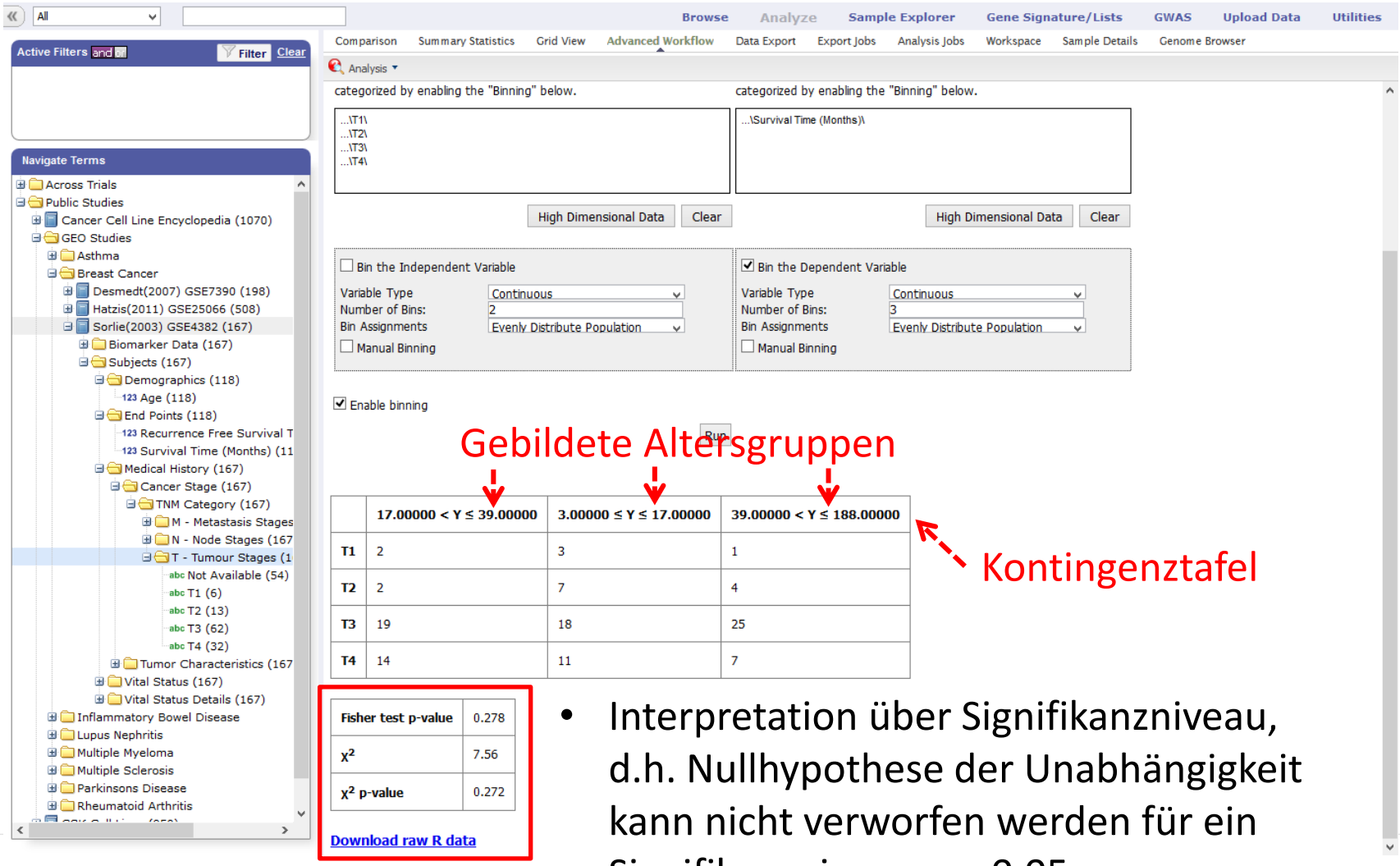
2. Überlebenszeit: das ist kein kategoriales Merkmal, aber tranSMART kann es dazu machen!



Ladies and Gentlemen, please start your engines!



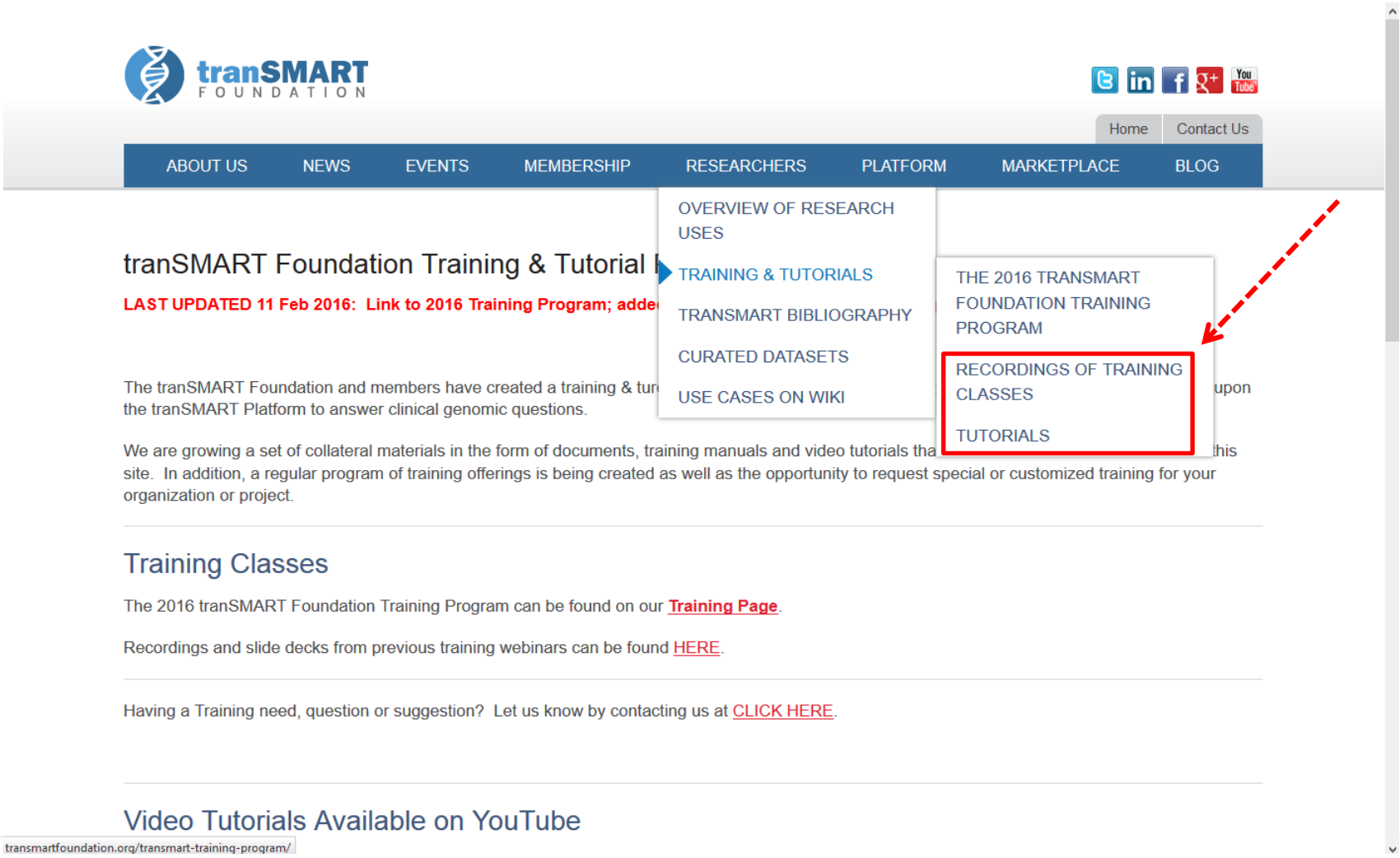
Ergebnis: p-Wert nach Fischer 0,278; p-Wert nach chi²-Test 0,272



	17.00000 < Y ≤ 39.00000	3.00000 ≤ Y ≤ 17.00000	39.00000 < Y ≤ 188.00000
T1	2	3	1
T2	2	7	4
T3	19	18	25
T4	14	11	7

- Interpretation über Signifikanzniveau, d.h. Nullhypothese der Unabhängigkeit kann nicht verworfen werden für ein Signifikanzniveau $\alpha = 0,05$

Webinars und Tutorials: umfangreiche Ressourcen verfügbar!



Vielen Dank!

Göttingen

Prof. Dr. Ulrich Sax
Ulrich.Sax@med.uni-goettingen.de
Christian Bauer
christian.bauer@med.uni-goettingen.de
Benjamin Baum
benjamin.baum@med.uni-goettingen.de

Institut für Medizinische Informatik, Universitätsmedizin Göttingen

Erlangen

Dr. Thomas Ganslandt
Thomas.Ganslandt@uk-erlangen.de
Christian Knell
christian.knell@fau.de
Jan Christoph
jan.christoph@fau.de

Lehrstuhl für Medizinische Informatik, Friedrich-Alexander-Universität Erlangen-Nürnberg

Leipzig

Matthias Löbe
matthias.loebe@imise.uni-leipzig.de
Sebastian Stäubert
sebastian.staeubert@imise.uni-leipzig.de

Institut für Medizinische Informatik, Statistik und Epidemiologie (IMISE), Universität Leipzig