Rejuvenated Medical Visualization

From Single-Patient to Cohort Study Data

Steffen Oeltze-Jafra

Tutorial Outline

Welcome

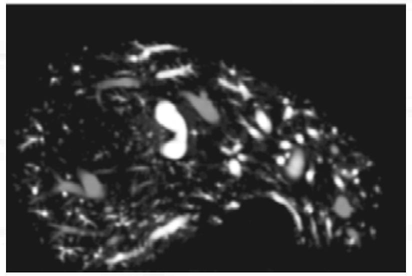
- Interactive Visualization of Whole-Body Medical Volume Data
  Ynnerman (40 min)
- From Static to Dynamic – Visualizing Real-Time Imaging Data
  Bruckner (40 min)

Coffee break (20min)

- From Anatomy to Physiology
  Hauser (35 min)
- From Single-Patient to Cohort Study Data
  Oeltze-Jafra (35min)
- Closing Words and Discussion
  All Presenters (30min)

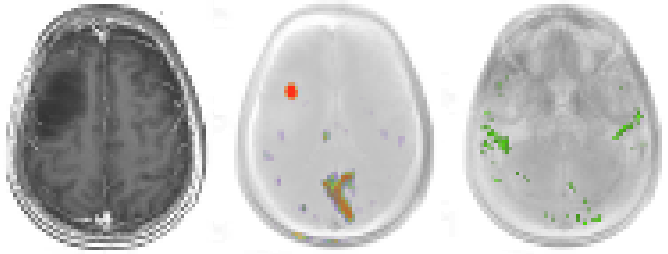
From Single-Patient to Cohort Study Data

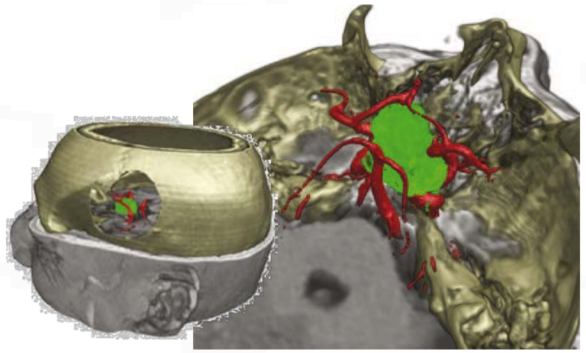
Single patient,
single body part



[Selle01]

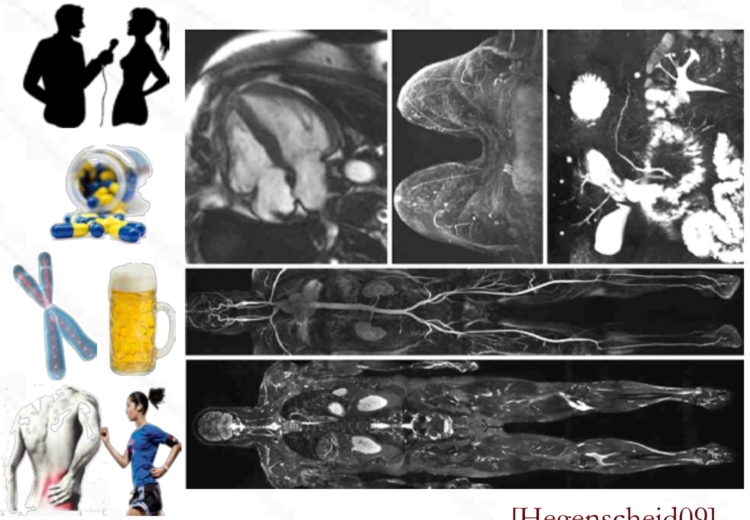
Single patient,
single body part,
multimodality or time-varying data





[Beyer07]

Thousands of subjects, longitudinal **non-image** (100s of variables) **and image data** (***whole-body*** and body parts, multi-modality, time-varying)



[Hegenscheid09]

?

Cohort Studies in Epidemiology

- Epidemiology investigates
  - Health-related events in defined populations (cohorts)
  - Factors contributing to disease outbreaks in individuals
  - Effectiveness of therapies and screening programs
- Investigation based on cohort studies
  - Often, carried out in waves over years (longitudinal)
  - Comprising thousands of individuals
  - Collecting hundreds of socio-demographic and biomedical variables for each individual

Modern Epidemiology

- Employs also *medical image data*, gene sequence data, blood and urine samples, ECG data as well as tissue samples
- These data are acquired efficiently, in large scale and high quality
- For ethical reasons, non-invasive imaging techniques are employed for largely healthy populations, e.g., MRI and US

Goals of Epidemiologic Research

1. Examination of *incidence* and *prevalence* of diseases
2. Identification of risk factors
3. Efficiency evaluation of preventive and therapeutic measures
4. Determination of differences between *healthy aging* and *pathologies in an early stage*

- No purely academic endeavor but has huge consequences
- Examples: passive smoking protection, safer sex campaigns, vaccination programs, cancer screening,…

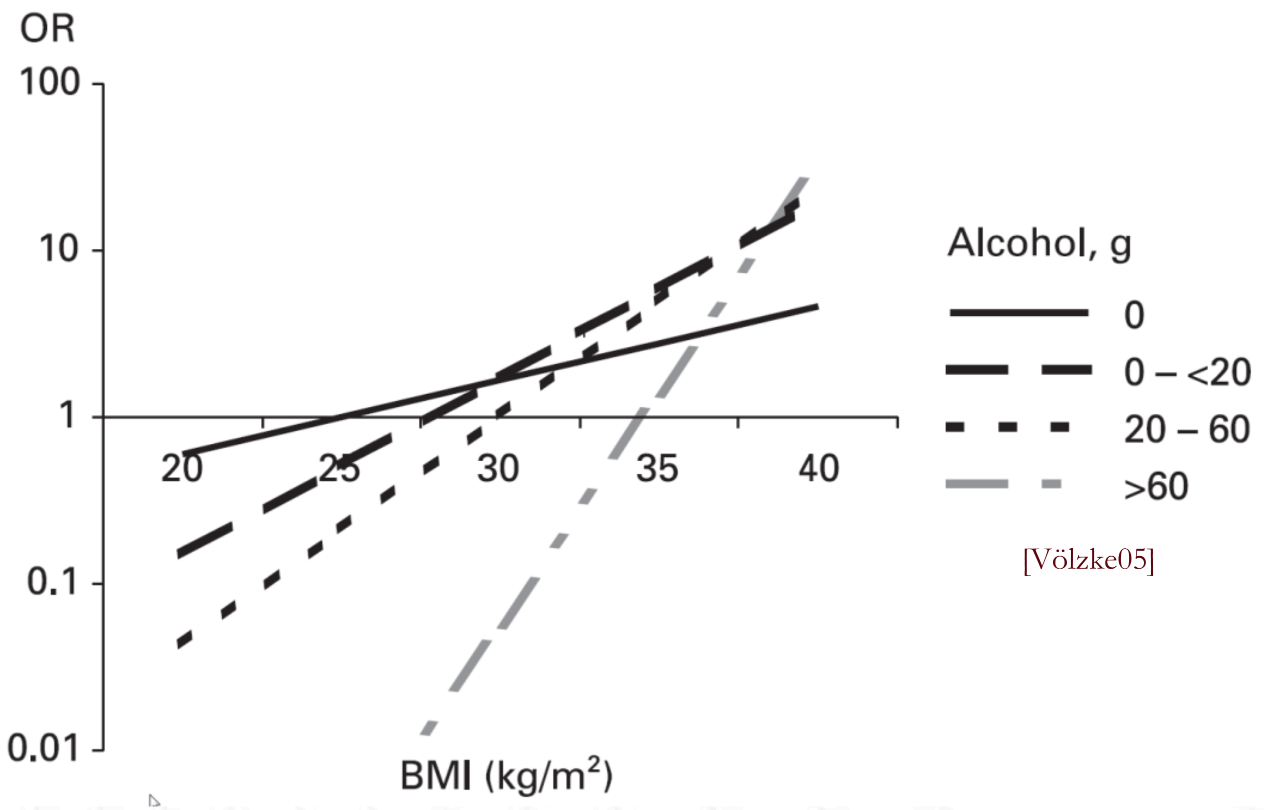
Some Results

- Breast cancer should not be operated in a radical manner in early stages of the disease.
- Radiation treatment may cause subsequent cancer diseases, in particular when children are treated. Limit the amount of radiation!
- Some risk factors depend linearly on a variable such as radiation dose or nicotine intake while others exhibit U- or J-shaped relations, e.g., the risk of mental disorders related to alcohol consumption or sleeping duration.

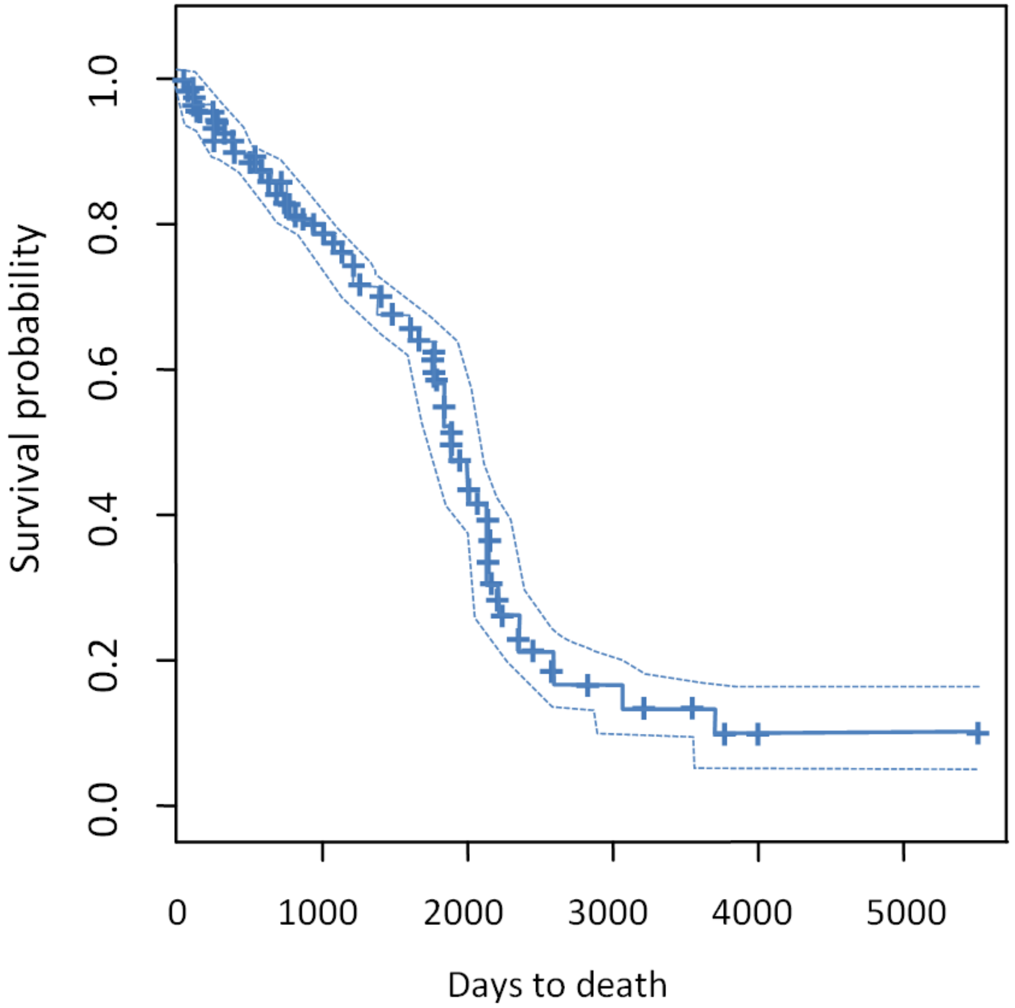
Data Characteristics

- Massive amount of heterogeneous data
- Different domains and aggregation levels
- (Different) variables per individual, organ, organ part
- Time-dependent and static variables
- Never perfectly reliable (consumption of alcohol, tobacco, …)
- Often sparse (treatment of diabetes),
- Often incomplete (not all people agree to the MRI exam)
- Dichotomous variables (presence of back pain)

Standard Visualizations



[Völzke05]



[Preim15]

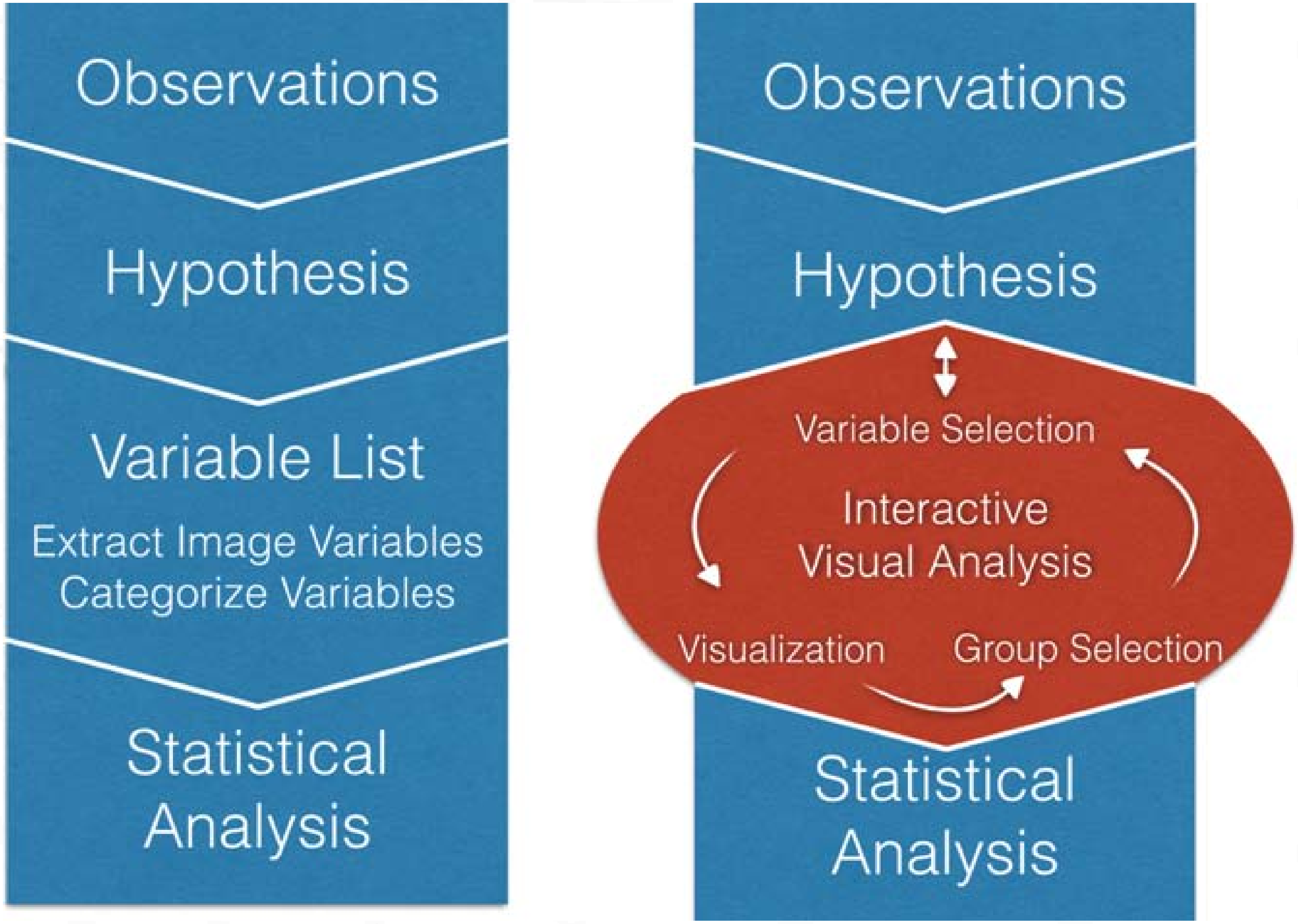
Left: *Interaction Term* describing the relative risk for the development of gall stones depending on body mass index (BMI) and alcohol consumption

Right: *Kaplan-Meier Curve* (including uncertainty) to characterize prognosis

Visual Analytics Approaches

Standard Analysis vs. Visual Analytics

- Old and new workflow in epidemiology
- Visual analytics allows for a more flexible and interactive selection of variables, a data-driven definition of groups and subgroups, and for the generation of new hypotheses



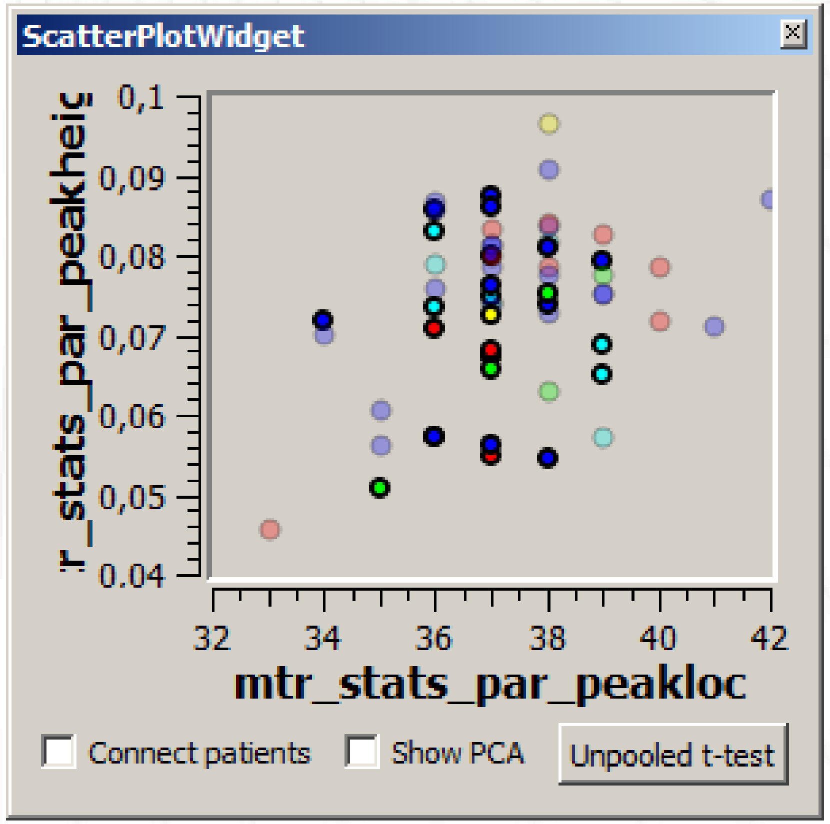
[Klemm14]

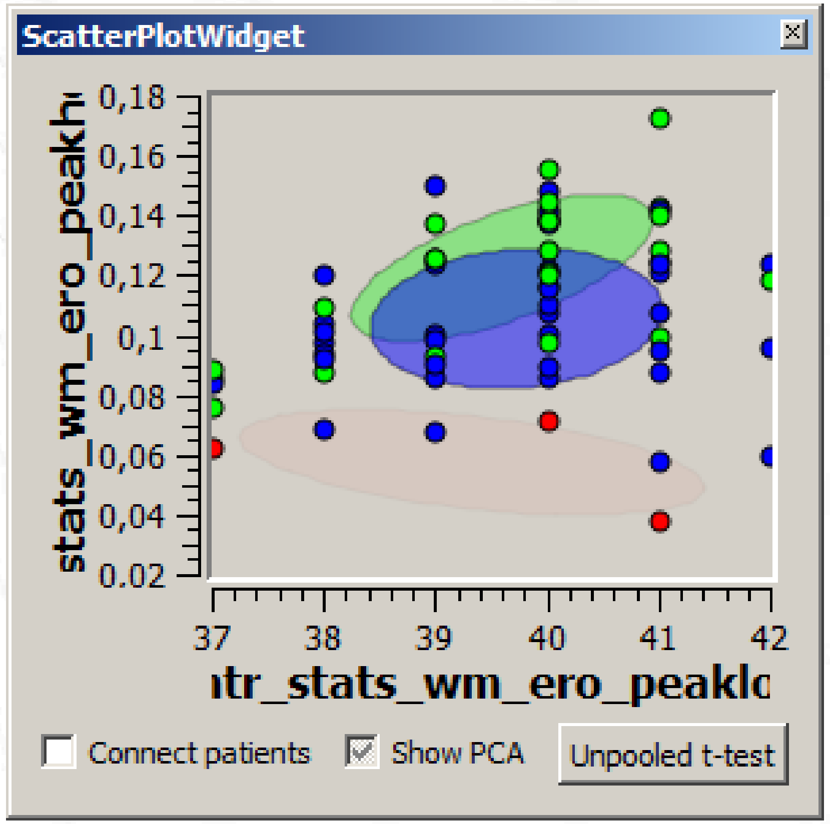
Potential of Visual Analytics

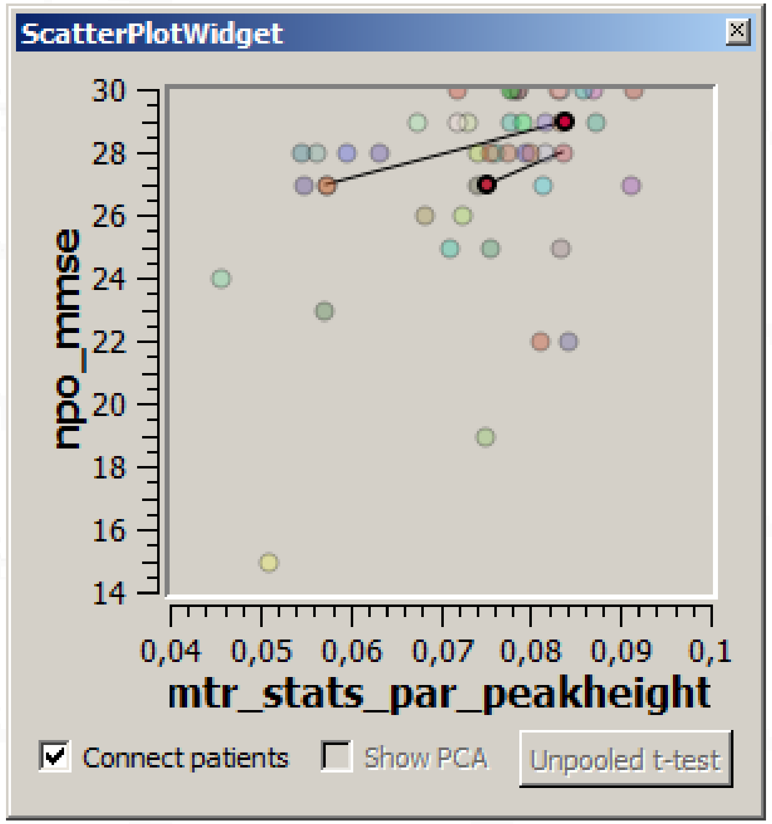
- Integration of data mining and interactive visual exploration
- **Data Mining**: machine learning techniques, such as pattern mining, classifiers, clustering, supervised learning techniques (Neural networks, support vector machines), dimension reduction
- **Interactive Visual Exploration**: visual queries, brushing and linking, multiple coordinated views, parallel coordinates, scatter plots, glyph-based displays, time-lines, graphs, and treemaps.
- **Integration**: analytics is not just a preprocess, but carefully integrated in the exploration, thus loops between visualization, exploration, and data mining are supported.

Examples of Visual Analytics Approaches

- Statistics is main research vehicle of epidemiologists
- Scatter plots with overlaid clusters and (local) regression lines



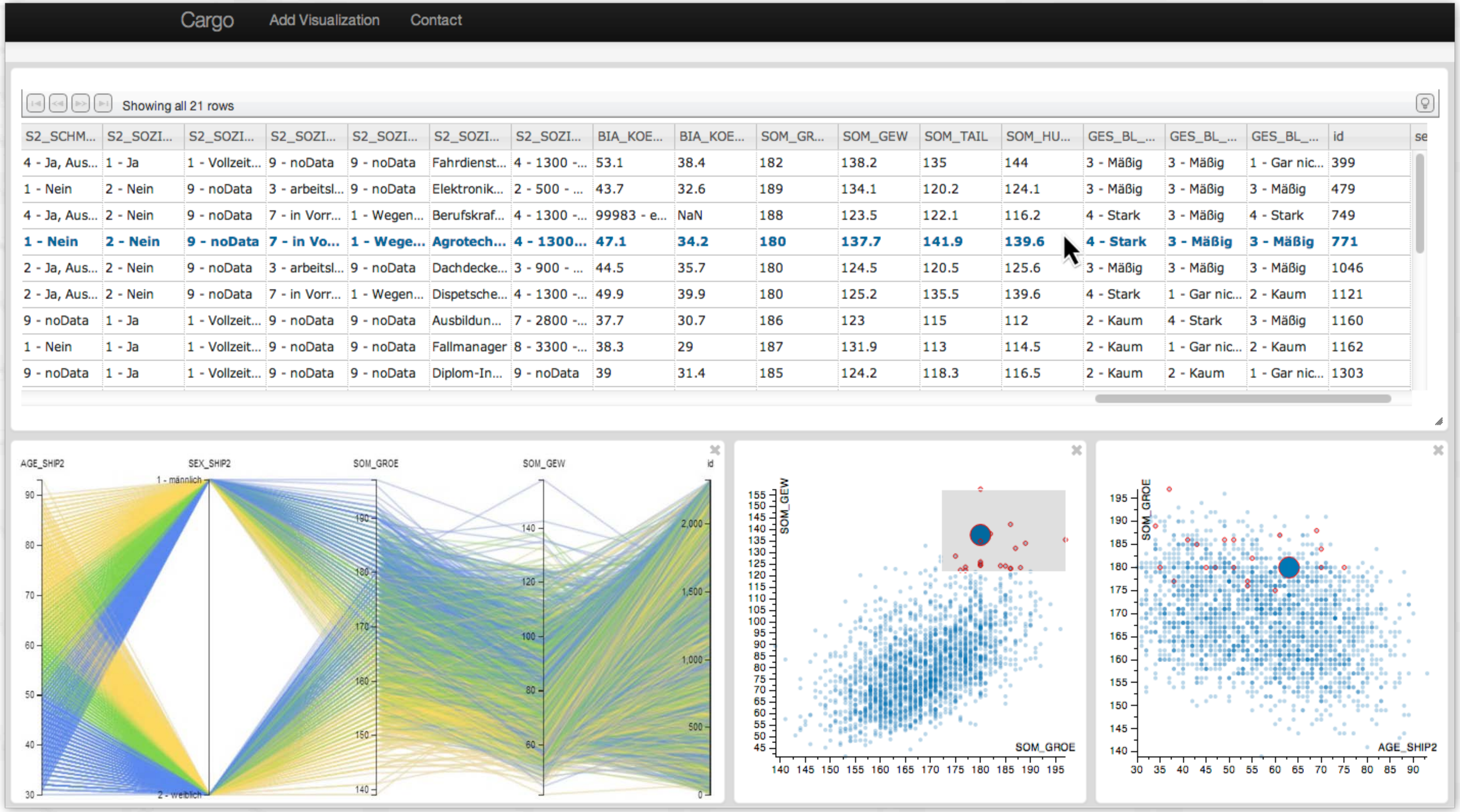




[Steenwijk10]

Examples of Visual Analytics Approaches

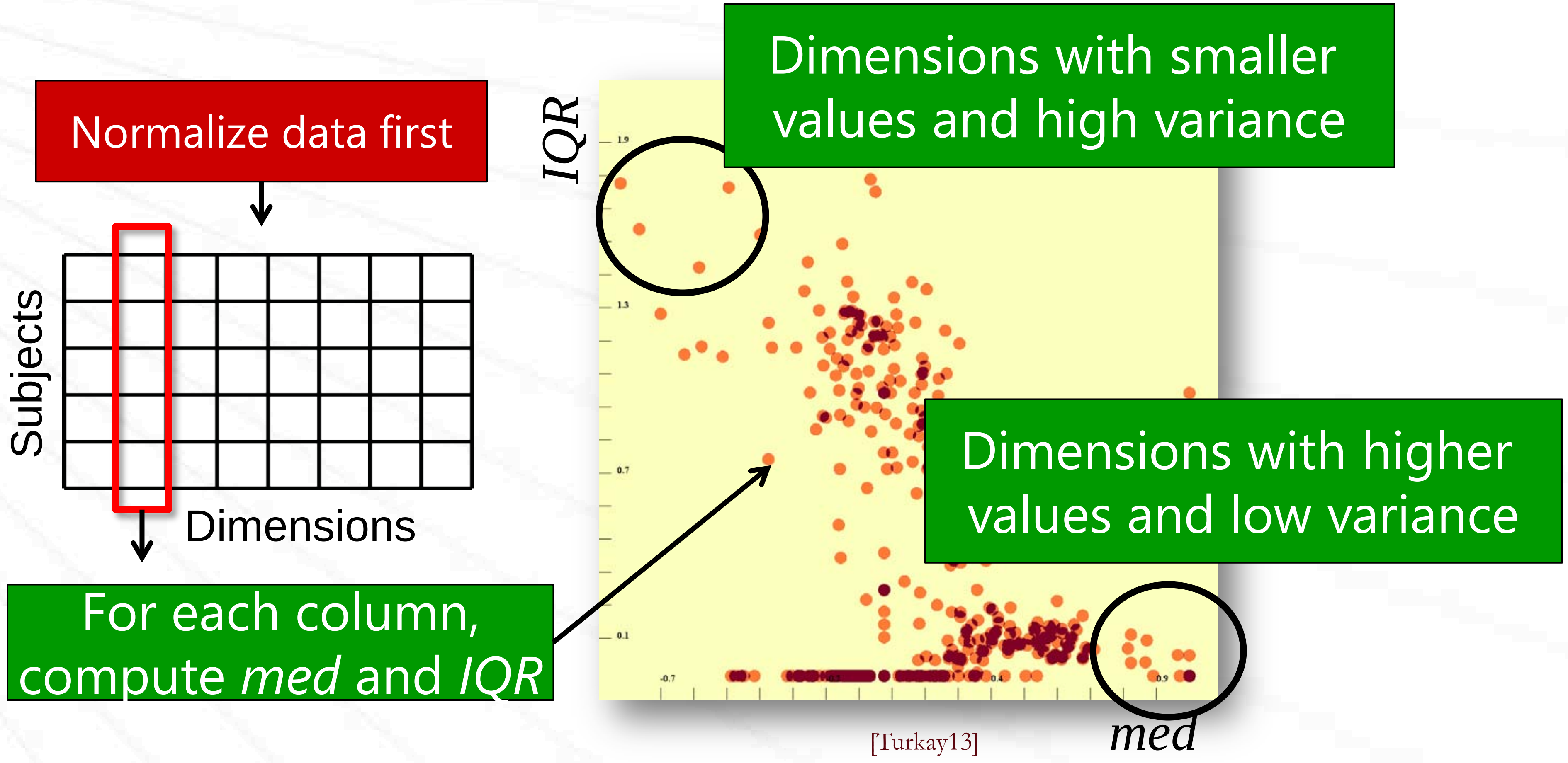
- Multiple Coordinated Views System
- Brushing & linking
- Clustering for data-driven definition of sub-groups



[Klemm14]

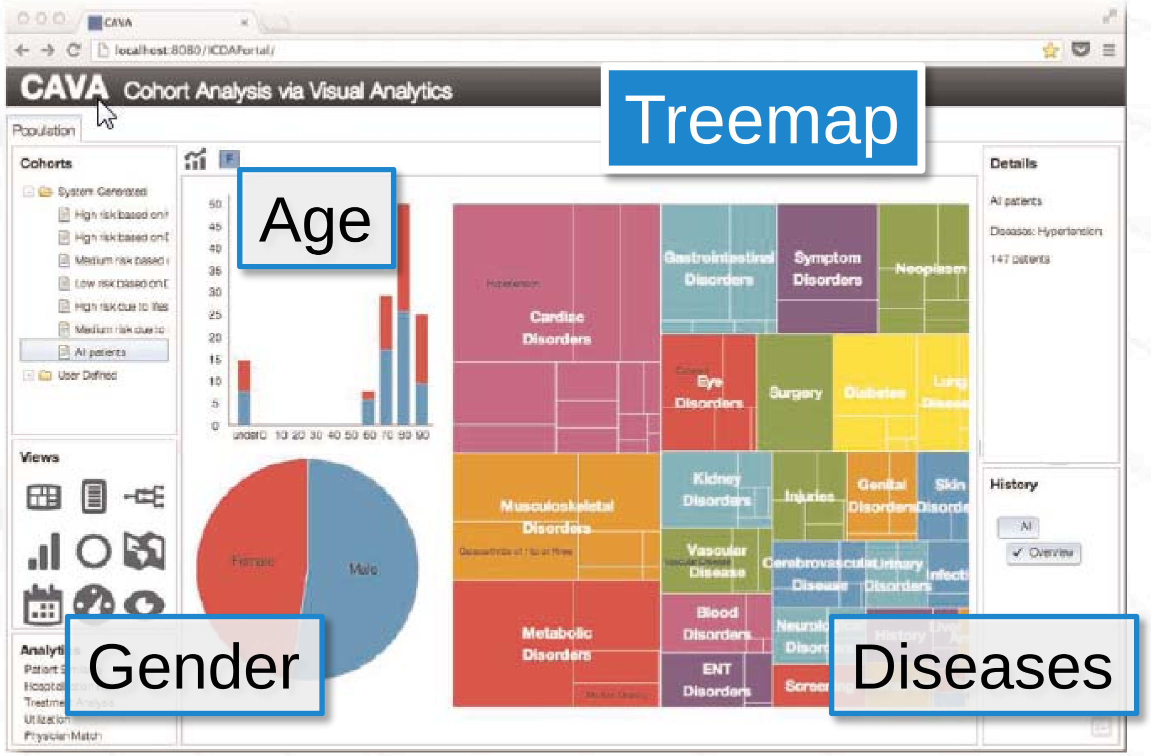
Examples of Visual Analytics Approaches

- Treat dimensions, NOT subjects, as first-order analysis objects [Turkay11]
- Compute statistics for each dimension, e.g., median, IQR, skewness, kurtosis
- Oppose statistics in plots and brush



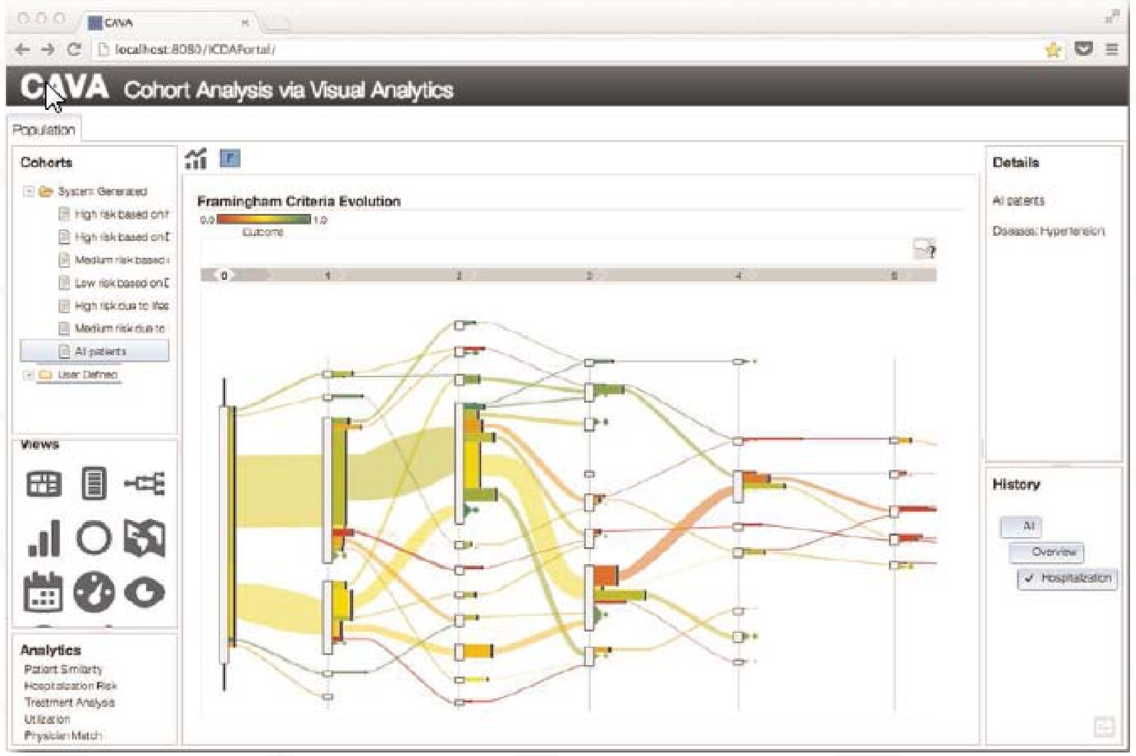
Examples of Visual Analytics Approaches

- Visualization techniques suitable for categorical data
- Time-oriented visualizations of events in longitudinal studies



[Zhang15]

Overview of a cohort

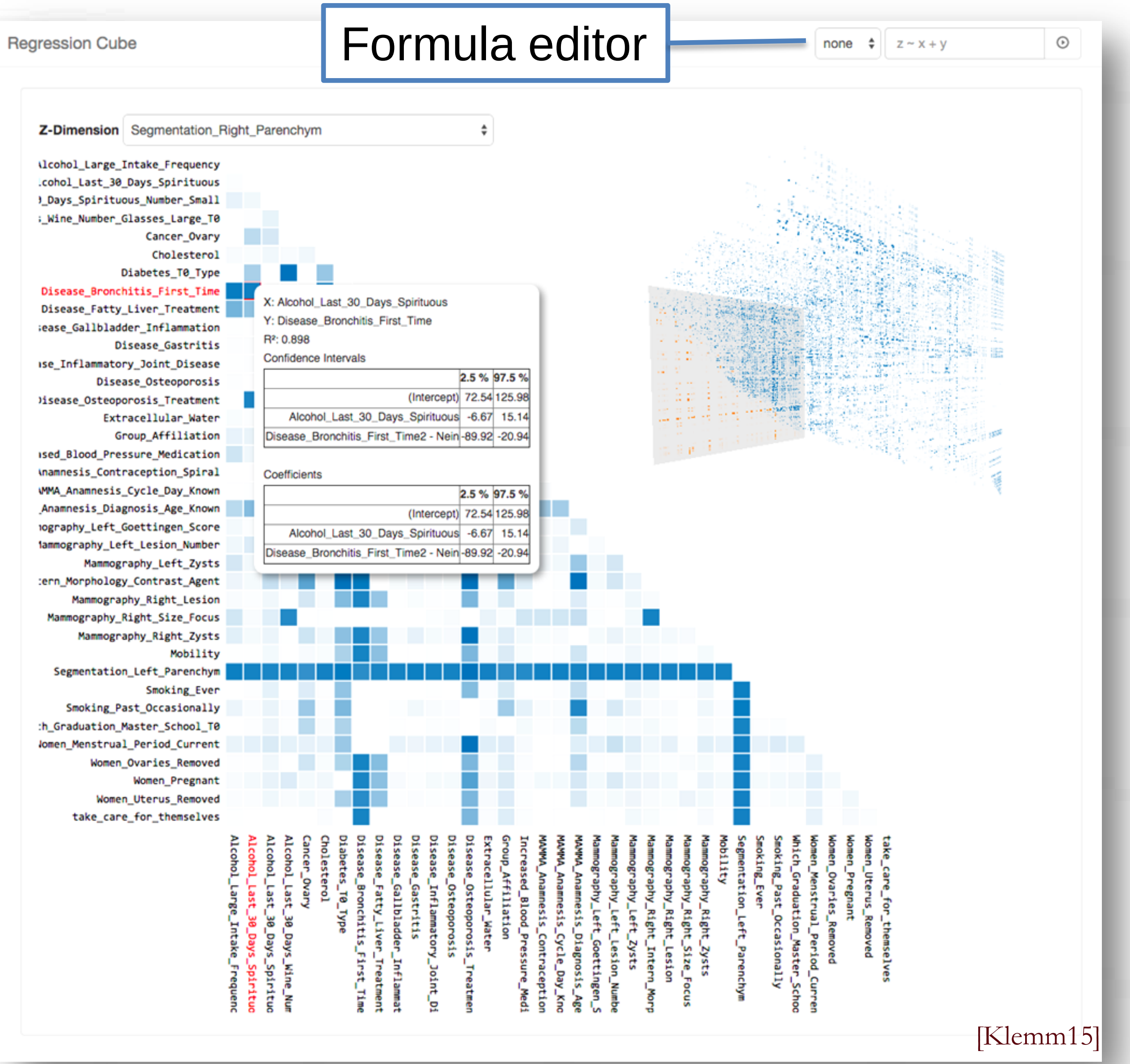


[Zhang15]

Symptom progression pathways

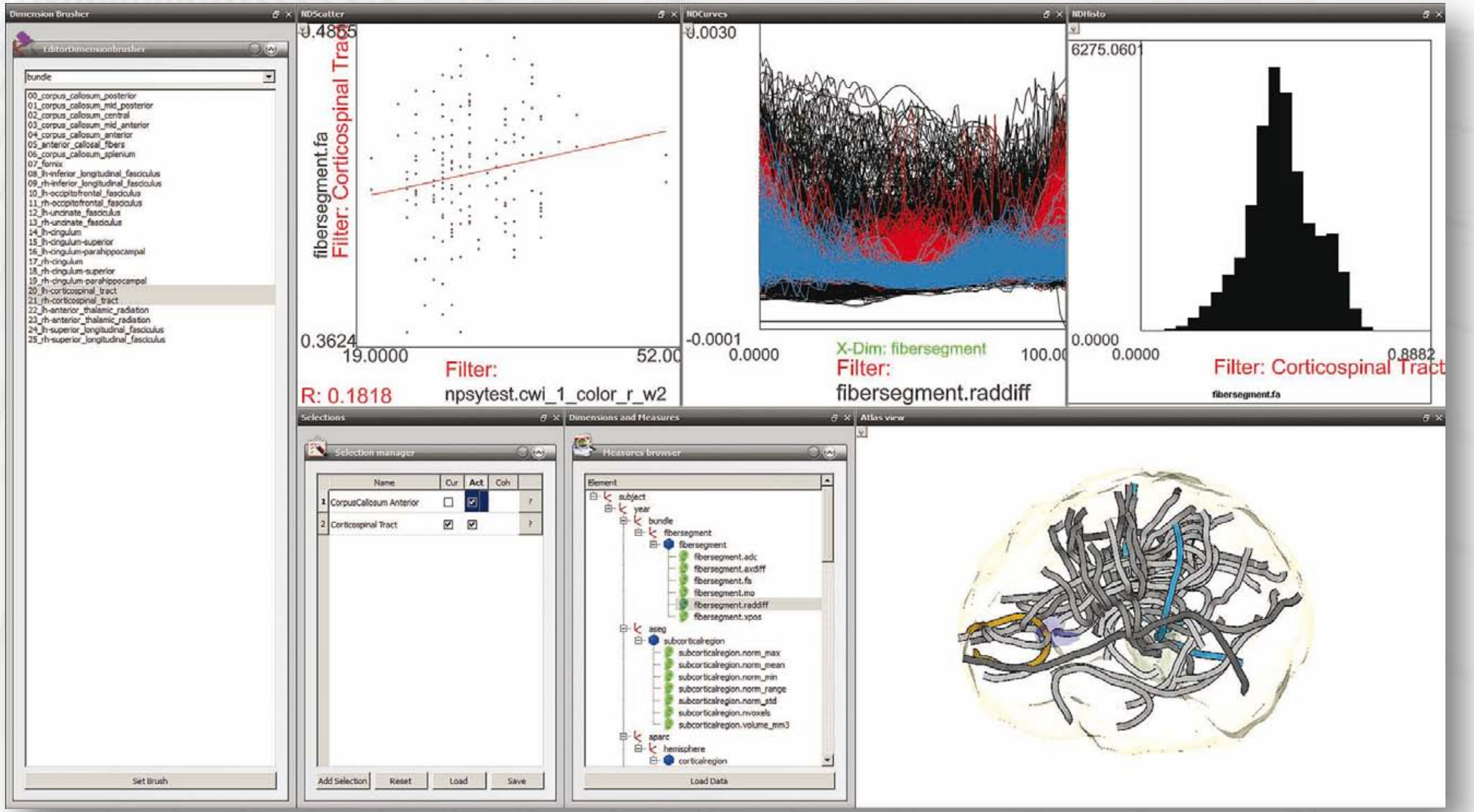
Regression-Based Visual Analytics

- **Key idea**: User specifies a target variable, usually a pathology, and combinations of (2 or 3) variables are checked w.r.t. a possible correlation (hypothesis-free)
- Regression cube for visualization
- Saturation encodes strength of correlation
- Come and see Paul's talk on Wed., 28th, VAST session at 8:30-10:10 in Room "Red"



[Klemm15]

Special Data Structures



[Angelelli14]

- Data cube model for seamless integration of entities with only partially overlapping dimensions
- Example: Fiber tracts and cortical regions with shared IDs of subject and study wave but disjoint sets of fractional anisotropy and cortical thickness measures
- Come and see Paolo's talk on Thurs., 29th, CG&A session at 14:00-15:40 in Room "Empire"

Visual Analytics: The Patent

- „Iterative Refinement of Cohorts Using Visual Exploration and Data Analytics“ (D. Gotz, A. Perer, Z. Zhang, IBM, granted 4/2014)
- **Preamble**: „A need exists for an integrated system that combines visual exploration and data analytics to ... refine cohorts ... and make new discoveries.“
- **Selected Claims** (out of 11):
  - Reduce cohort using one or more visual filters,
  - Visualize selected cohort using a selected view,
  - Expand current cohort by one or more analytical methods,
  - Modify current cohort using one or more additional reductions and extensions

Visual Analytics Prototyping

- Prototyping using modern web technologies allows for a fast feedback loop between computer scientists and epidemiologists
- Client-Server structure to outsource heavy computations on a server machine, keeping system requirements on the client low

[Klemm14]

Image-Centric Cohort Studies

Potential of Image Data

- Depiction of the vital organs of a large population
- Representation of the broad variability in anatomy and physiology
- Characterization of health and disease
- Differentiation between effects of normal aging and pathologies

Disclaimer:

The individual is of less interest to epidemiologists, just like the realistic, high-quality visualization of its specific image data.

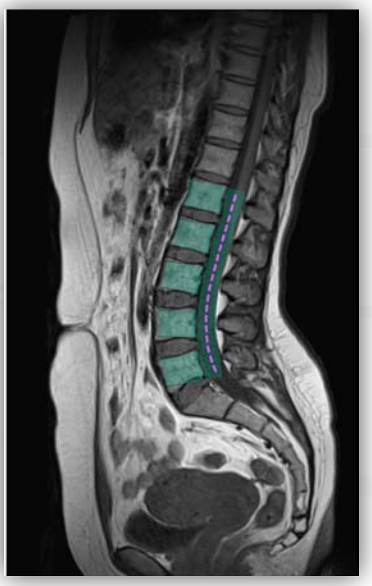
Data Acquisition

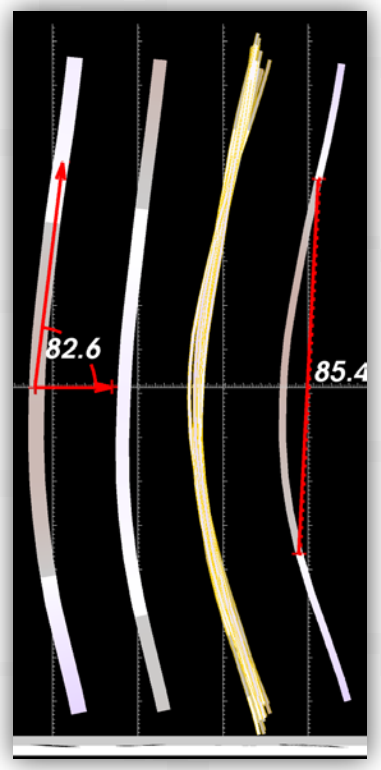
- All *epidemiological instruments* need to be applied in a highly standardized manner → minimize intra-observer variability
- Scanners, sequences and protocols are highly standardized
- No updates are allowed in the whole (>5 year!) cycle
- Ethical and economic aspects are essential
  - Primarily Ultrasound and MRI

Problem: not all people may agree to acquisition or may be eligible for it, e.g., pacemaker patients in MRI

Data Processing

- Image data itself are too complex
- Epidemiologists are interested in derived measures related to:
    - Shape, Size, Angle,
    - Location, neighborhood,
    - Tissue density,…
- Huge amount of data requires segmentation and quantification to be as automatic and reproducible as possible





Example Studies

- SHiP (Study of Health in Pomerania), Longitudinal study in 3 waves, wide range of diseases, focus: diabetes, cardio vascular diseases, diseases of the bile duct and liver
- Rotterdam Study (Focus: elderly persons (>55 years) and cardio vascular diseases, Start: 1990, >900 PubMed publications)
- UK Biobank (500.000 persons, 2006-2010)
- Cognitive Aging Study (Bergen, Norway) [Ystad09]
- Nationale Kohorte (Germany, Start: 2014) [GNC14]
  - 200.000 test persons (aged 20-69, 20% intensified level 2 examination)
  - ~30.000 MRI data will be acquired in 5 centers
  - Automatic quality control and normalization as essential aspects (MR scanners do not have perfectly the same characteristics)

Study of Health in Pomerania (SHiP)

Study in 3 waves [Völzke11]

- SHiP-0: 1997-2001, n=4308 (f/m), age: 20-79
- SHiP-1: 5-year-follow up, n=3300
- SHiP-2: 12-year follow-up, n=2500
  (less participants as always; drop out)
  SHiP-Trend: new sample drawn, n=8016

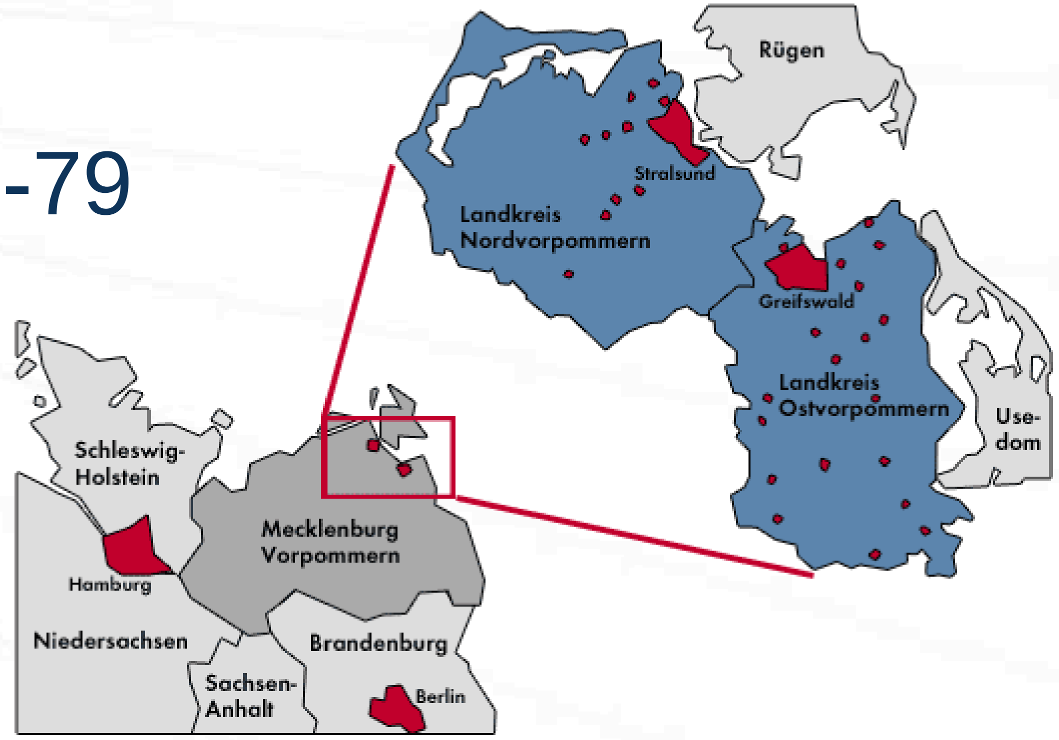


Image Data [Hegenscheid09]

- Whole-body MRI: 1.5 Tesla, basis program + contrast-enhanced cardio MRI + MRA for men and MR Mammography for women (up to 40 datasets per study participant)

Case Study: Analysis of Lower Back Pain

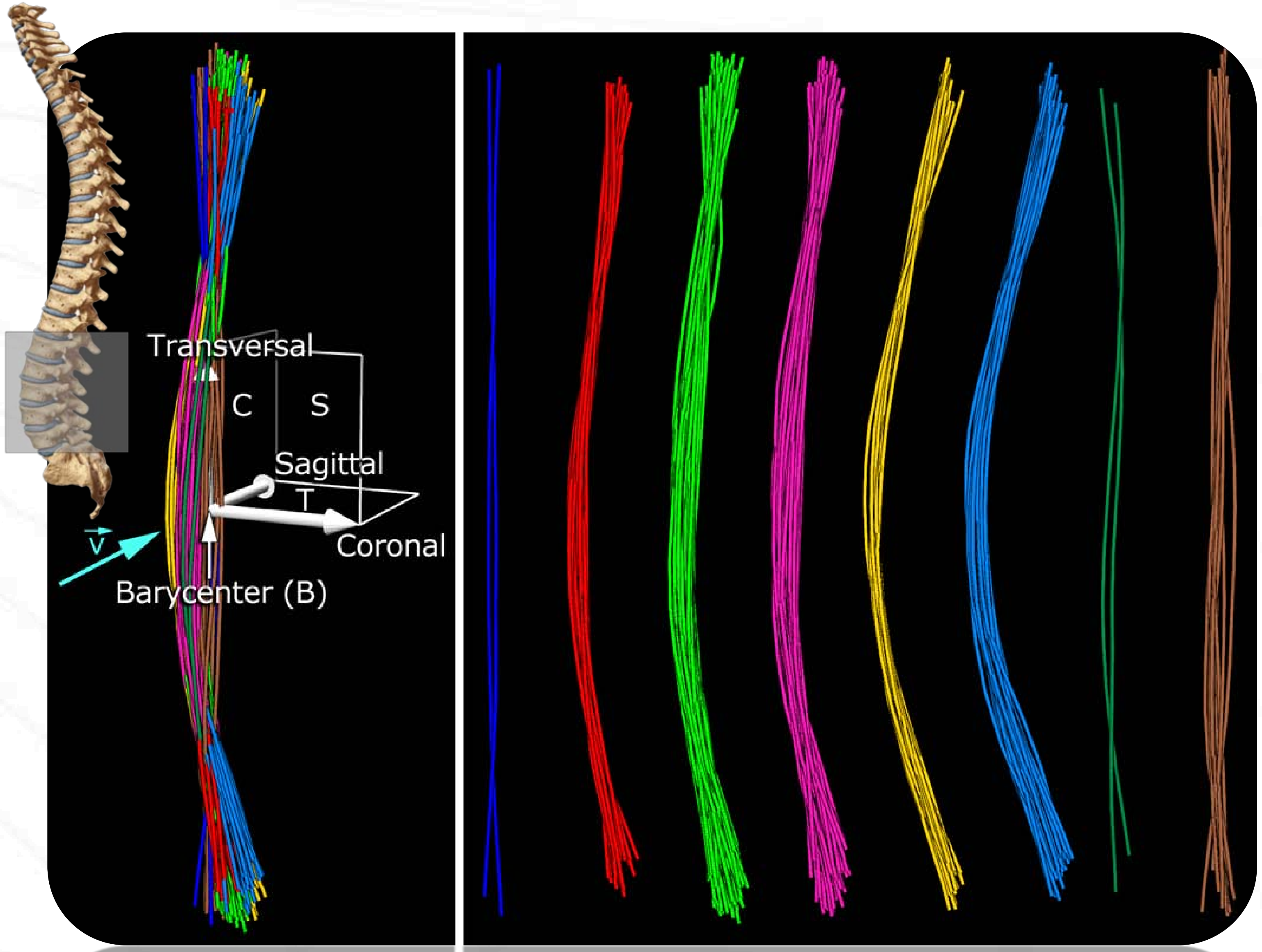
- 129 variables (only 15 scalar) of SHiP are possibly related to back pain
- Image data: T1/T2-weighted MRI data, rather low res (1.1x1.1x4 mm)
- Known correlations:
  - Back pain is related to age, depression, physical stress, smoker status [Manek05]
  - Curvature of spine is related to age, body height, weight
- Visual Analytics goals:
  - Confirm known or assumed correlations
  - Identify new correlations
  - Differentiate between pathological and age-related deformations of the spine

Case Study: Analysis of Lower Back Pain

- Task: robust, automatic detection of the spine
- Hierarchical Finite Element method applied to >2000 patients [Rak13]
- Detection combined with centerline extraction (repr. by 100 vertices)
- Successful in ~80 percent

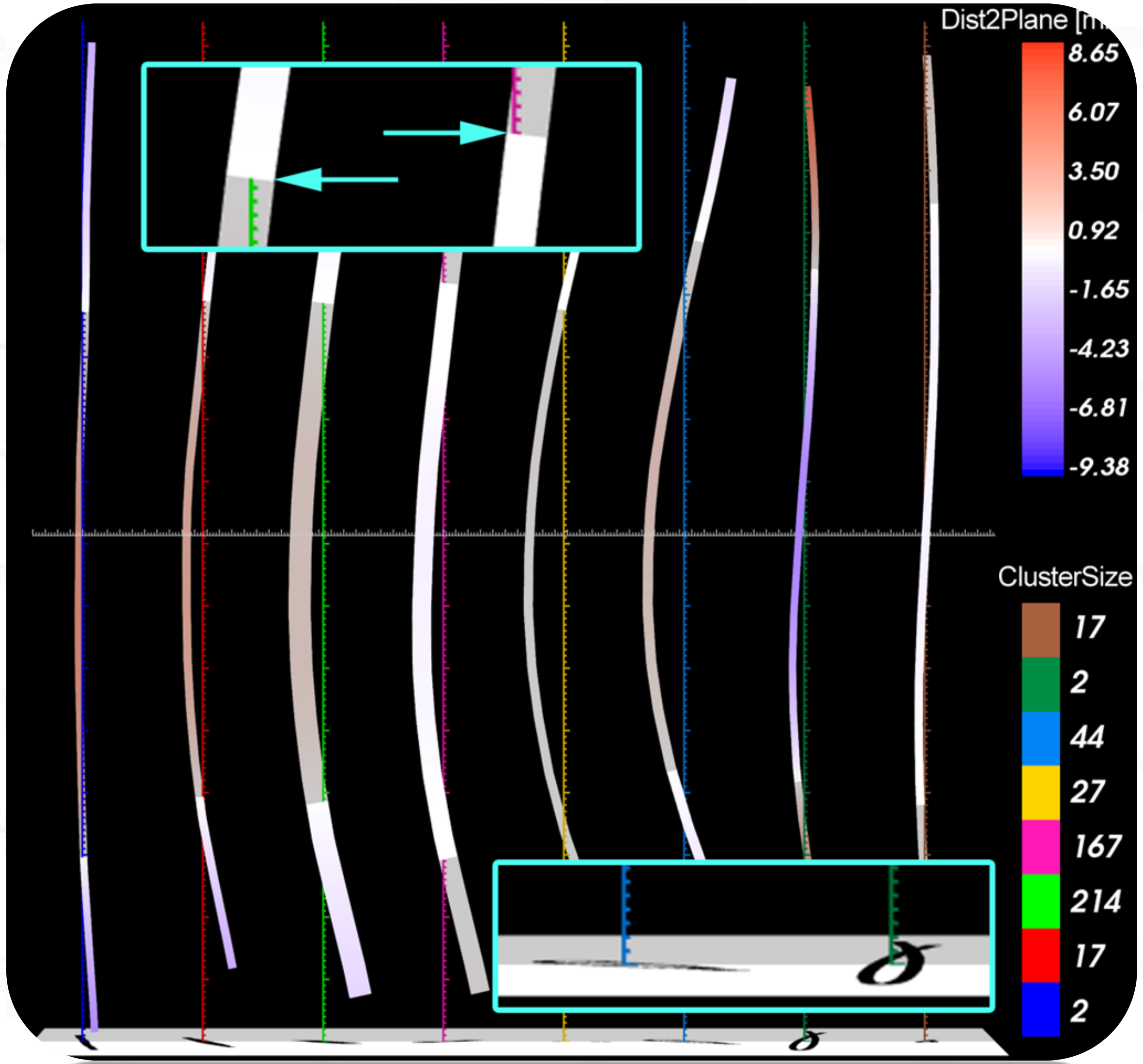
Case Study: Analysis of Lower Back Pain

- Task: identification of sub-groups of participants with similar spine canal bending
- Agglomerative Hierarchical Clustering approach [Klemm13]
- Automatic determination of number of clusters [Salvador04]



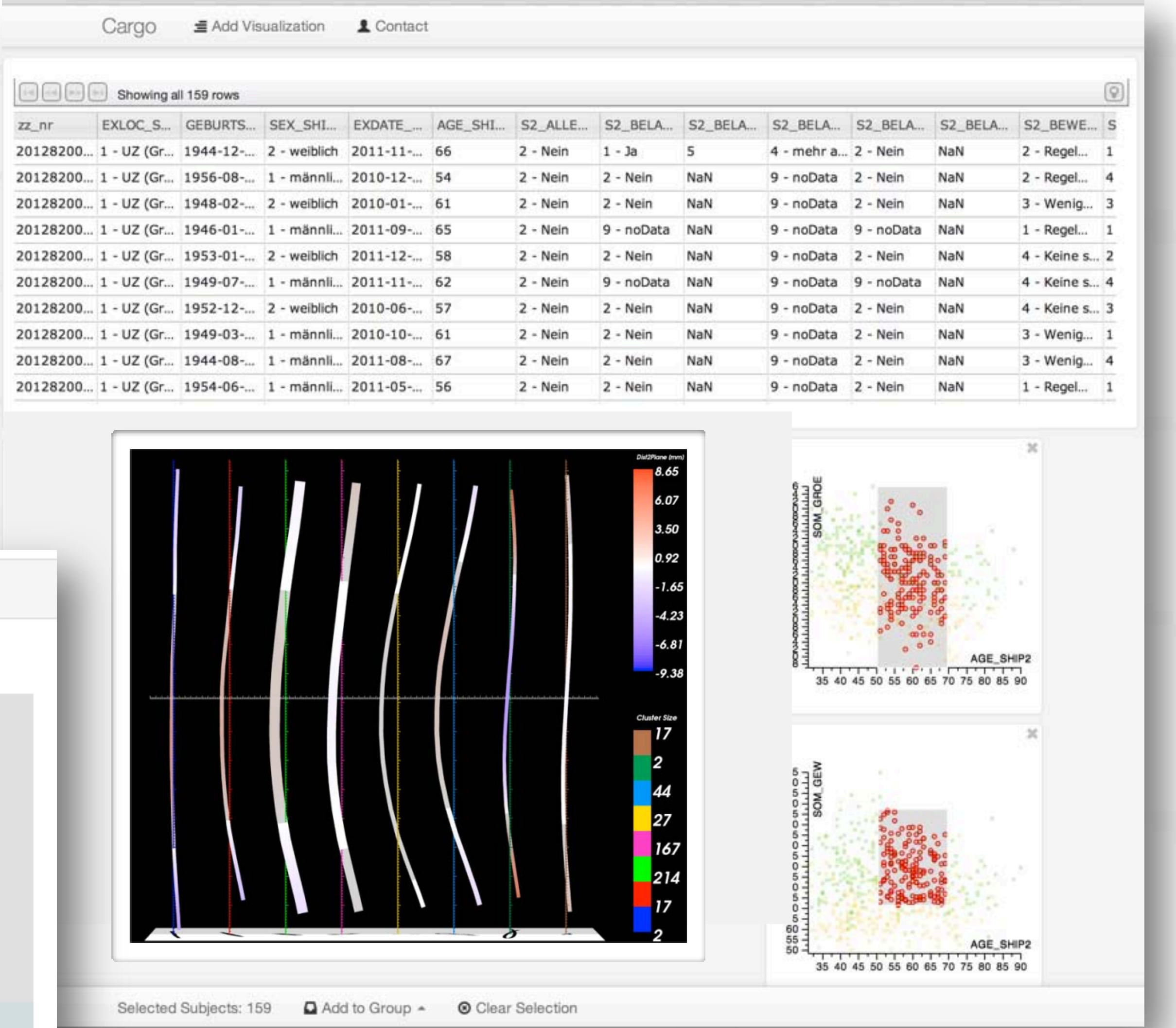
Case Study: Analysis of Lower Back Pain

- Task: display of clustering results
- Ribbon-based visualization of cluster representatives
- Width encodes cluster size
- Color encodes distance to midsagittal plane
- Shadows on projection plane indicate bending strength orthogonal to midsagittal plane

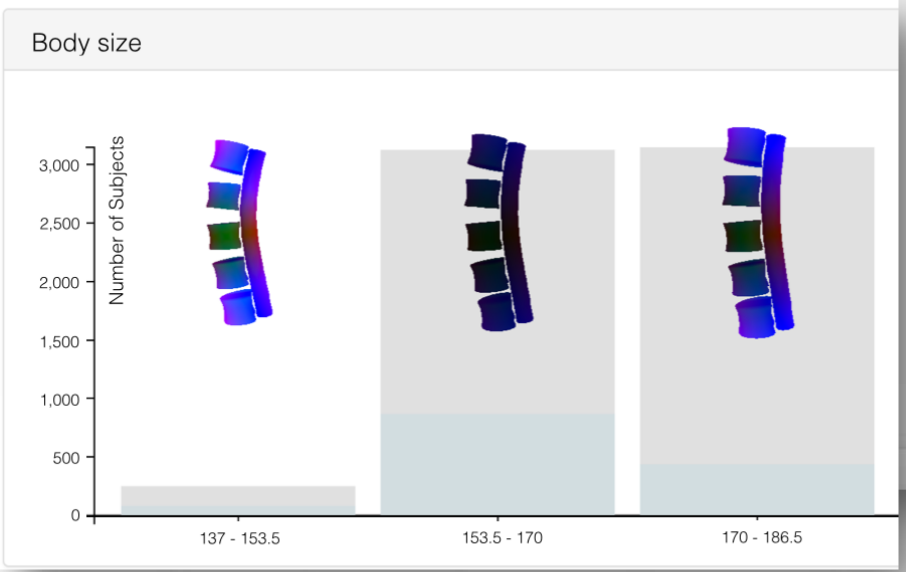


Case Study: Analysis of Lower Back Pain

- Tasks: relation of clusters to cohort variables and search for correlations
- Integration of 3D view into framework
- Bidirectional linking between 3D view and all other views
- Augmentation of InfoVis plots by mean shapes of lumbar spine



zz_nr	EXLOC_S...	GEBURTS...	SEX_SHI...	EXDATE_...	AGE_SHI...	S2_ALLE...	S2_BELA...	S2_BELA...	S2_BELA...	S2_BELA...	S2_BELA...	S2_BEWE...	S
20128200...	1 - UZ (Gr...	1944-12-...	2 - weiblich	2011-11-...	66	2 - Nein	1 - Ja	5	4 - mehr a...	2 - Nein	NaN	2 - Regel...	1
20128200...	1 - UZ (Gr...	1956-08-...	1 - männli...	2010-12-...	54	2 - Nein	2 - Nein	NaN	9 - noData	2 - Nein	NaN	2 - Regel...	4
20128200...	1 - UZ (Gr...	1948-02-...	2 - weiblich	2010-01-...	61	2 - Nein	2 - Nein	NaN	9 - noData	2 - Nein	NaN	3 - Wenig...	3
20128200...	1 - UZ (Gr...	1946-01-...	1 - männli...	2011-09-...	65	2 - Nein	9 - noData	NaN	9 - noData	9 - noData	NaN	1 - Regel...	1
20128200...	1 - UZ (Gr...	1953-01-...	2 - weiblich	2011-12-...	58	2 - Nein	2 - Nein	NaN	9 - noData	2 - Nein	NaN	4 - Keine s...	2
20128200...	1 - UZ (Gr...	1949-07-...	1 - männli...	2011-11-...	62	2 - Nein	9 - noData	NaN	9 - noData	9 - noData	NaN	4 - Keine s...	4
20128200...	1 - UZ (Gr...	1952-12-...	2 - weiblich	2010-06-...	57	2 - Nein	2 - Nein	NaN	9 - noData	2 - Nein	NaN	4 - Keine s...	3
20128200...	1 - UZ (Gr...	1949-03-...	1 - männli...	2010-10-...	61	2 - Nein	2 - Nein	NaN	9 - noData	2 - Nein	NaN	3 - Wenig...	1
20128200...	1 - UZ (Gr...	1944-08-...	1 - männli...	2011-08-...	67	2 - Nein	2 - Nein	NaN	9 - noData	2 - Nein	NaN	3 - Wenig...	4
20128200...	1 - UZ (Gr...	1954-06-...	1 - männli...	2011-05-...	56	2 - Nein	2 - Nein	NaN	9 - noData	2 - Nein	NaN	1 - Regel...	1



Case Study: Analysis of Lower Back Pain

- Analysis revealed correlations of lower back pain with
  - leg fatigue,
  - physically heavy work,
  - body weight,
  - dyspnoea, and
  - headache intensity (particularly interesting)
- Analysis, so far, is restricted to the overall spine shape
- Further insights are expected when the following is considered:
  - The shape and size of individual vertebrae
  - The angle between the central spine axis and the vertebrae
  - The distance between individual vertebrae

Concluding Remarks

- Visual analytics approach is subject to discussions, e.g., the value of clustering, subspace-clustering? Parameter sensitivity?
- *Hypothesis generation* using *interactive visual* analysis of the *entire* data pool is a new concept to epidemiologists
- Integration of statistics and statistical graphics is essential to increase acceptance of new concepts, tools, and views
- When working with epidemiologists there is no direct influence on patient care but potentially a large indirect influence by the refinement of guidelines for diagnosis and treatment processes.

Outlook

- Pair analytics
  - Expert and computer scientist cooperatively analyze epidemiological data
  - Localized or distributed (analysis performed at one or separate locations)
  - Consequence: one workspace or two separate workspaces, each with input devices, requiring synchronization and an appropriate GUI-design

- Uncertainty-aware visual analytics
  - Ambiguous subgroup definition yields uncertain statistics
  - Measurement inaccuracies, e.g. of blood pressure, or unreliable self reports, e.g. on eating and drinking habits, yield uncertain variables
  - Consequence: model uncertainties, consider them in the analysis, and convey them in the visualization

Acknowledgements

SHiP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grant no. 03ZIK012), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Whole-body MR imaging was supported by a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-Vorpommern. The work on the analysis of lower back pain was supported by the DFG Priority Program 1335: Scalable Visual Analytics.

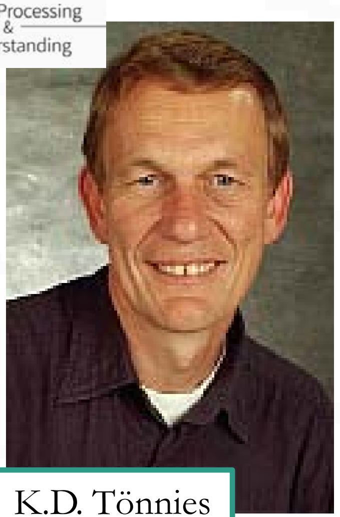

P. Klemm


B. Preim


H. Völzke


K. Hegenscheid


M. Rak


K.D. Tönnies

Literature

[Angelelli14] Angelelli P., Oeltze S., Turkay C., Haasz J., Hodneland , Lundervold A., Lundervold A., Preim B., Hauser H., Interactive Visual Analysis of Heterogeneous Cohort Study Data, *IEEE CG&A*, 34(5):70-82, 2014.

[Beyer07] J. Beyer, M. Hadwiger, S. Wolfsberger, K. Bühler. High-Quality Multimodal Volume Rendering for Preoperative Planning of Neurosurgical Interventions*, IEEE TVCG*, 13(6):1696-1703, 2007.

[GNC14] German National Cohort (GNC) Consortium. The German National Cohort: aims, study design and organization", *Eur. J. of Epidemiology*, 29: 371-382, 2014.

[Hegenscheid09] K. Hegenscheid, J. P. Kühn, H. Völzke, R. Biffar, N. Hosten, R. Puls, Whole-body magnetic resonance imaging of healthy volunteers: pilot study results from the population-based SHIP study. *RöFo*, 181(8):748-759, 2009.

[Klemm15] P. Klemm, K. Lawonn, S. Glaßer, U. Niemann, K. Hegenscheid, H. Völzke, B. Preim. 3D Regression Heat Map Analysis of Population Study Data, *IEEE TVCG*, in print, 2015.

[Klemm14] P. Klemm, S. Oeltze-Jafra, K. Lawonn, K. Hegenscheid, H. Völzke, B. Preim. Interactive Visual Analysis of Image-Centric Cohort Study Data, *IEEE TVCG*, 20(12):1673-1682, 2014.

[Klemm13] P. Klemm, K. Lawonn, M. Rak, B. Preim, K.D. Tönnies, K. Hegenscheid, H. Völzke, S. Oeltze. Visualization and Analysis of Lumbar Spine Canal Variability in Cohort Study Data, Proc. of VMV, pp. 121-128, 2013.

[Manek05] N. J. Manek, A. J. MacGregor. Epidemiology of back disorders: prevalence, risk factors, and prognosis. *Current Opinion in Rheumatology*, 17(2):134–140, 2005.

[Preim15] B. Preim, P. Klemm, H. Hauser, K. Hegenscheid, S. Oeltze, K. Toennies, H. Völzke. Visual Analytics of Image-Centric Cohort Studies in Epidemiology, *Visualization in Medicine and Life Sciences*, in print, 2015

Literature

[Rak13] M. Rak, K. Engel, K.D. Tönnies. Closed-Form Hierarchical Finite Element Models for Part-Based Object Detection, *Proc. of VMV*, pp. 137-144, 2013.

[Salvador04] S. Salvador, P. Chan. Determining the Number of Clusters/Segments in Hierarchical Clustering/Segmentation Algorithms. *Proc. of Tools with Artificial Intelligence (ICTAI)*, pp. 576 – 584, 2004.

[Selle01] H. Hahn, B. Preim, D. Selle, H.-O. Peitgen. Visualization and Interaction Techniques for the Exploration of Vascular Structures, *Proc. of IEEE Visualization*, pp. 395-402, 2001.

[Steenwijk10] M. D. Steenwijk, J. Milles, M. A. Buchem, J. H. Reiber, C. P. Botha. Integrated visual analysis for heterogeneous datasets in cohort studies. *Proc. of IEEE VisWeek Workshop on Visual Analytics in Health Care*, 2010.

[Turkay13] C. Turkay, A. Lundervold, A.J. Lundervold, H. Hauser. Hypothesis Generation by Interactive Visual Exploration of Heterogeneous Medical Data. *Proc of. CHI-KDD* (1-12), 2013.

[Turkay11] C. Turkay, P. Filzmoser, H. Hauser. Brushing Dimensions-A Dual Visual Analysis Model for High-Dimensional Data. *IEEE TVCG*, 17(12): 2591-2599, 2011.

[Völzke11] H. Völzke, D. Alte, C.O. Schmidt et al.: Cohort Profile: The Study of Health in Pomerania. *International Journal of Epidemiology*, 40(2): 294-307, 2011.

[Völzke05] H. Völzke, S.E. Baumeister, D. Alte, W. Hoffmann, C. Schwahn, P. Simon, U. John, M.M. Lerch: Independent Risk Factors for Gallstone Formation in a Region with High Cholelithiasis Prevalence. *Digestion*, 71(2) :97–105. (2005)

[Ystad09] M. Ystad, A. Lundervold, E.Wehling, T. Espeseth, H. Rootwelt, L.Westlye, M. Andersson, S. Adolfsdottir, J. Geitung, A. Fjell, et al. Hippocampal volumes are important predictors for memory function in elderly women. *BMC medical imaging*, 9(1):17, 2009.

[Zhang15] Z. Zhang, D. Gotz, A. Perer. Iterative cohort analysis and exploration, *Information Visualization*, 14(4):289-307, 2015.