Knowledge Discovery from Cohorts, Electronic Health Records and further Patient-related data

Myra Spiliopoulou

KDD 2018, London, August 19

Cohorts

[Glenn, 2005]

The term "cohort"

Quoting [Glenn, 2005], page 2: "The term *cohort* originally referred to a group of warriors or soldiers, and the term is now sometimes used in a very general sense to refer to a number of individuals who have some characteristic in common."

The term "cohort" in "cohort analysis"

Quoting [Glenn, 2005], page 2: "Here and in other literature on cohort analysis, however, the term is used in a more restricted sense to refer to those individuals (human or otherwise) who experienced a particular event during a specified period of time. The kind of cohort most often studied by social scientists is the human *birth cohort*, that is, those persons born during a given year, decade, or other period of time."

Cohort Analysis [Glenn, 2005]

The term “cohort analysis”

Quoting [Glenn, 2005], page 3: “The term *cohort analysis* is usually reserved for studies in which two or more cohorts are compared with regard to at least one dependent variable measured at two or more points in time.”

Purposes of Cohort Analysis [Glenn, 2005], pages 1-2

- “Assessing the effects of aging”
- “Understand[ing] the sources and nature of social, cultural and political change.”

Counter-examples – [Glenn, 2005], page 3

- *Cross-sectional study*: Comparison of different groups of individuals with respect to some characteristic/variable – such a study “is conducted with data collected at one point in time, or, more accurately, within a short period of time.”
- *Panel study:* Comparison of the the attitudes of a group of individuals at two distinct timepoints – such a study “measures the characteristics of the same individuals at more than one point in time.”

Learning from Mobile Data - Outline

- Mobile sensing and crowdsensing
- Mobile sensing for mental health
- Mobile sensing and Ecological Momentary Assessments

Mobile Crowdsending and Computing [Guo et al., 2015]

Mobile Crowdsensing and Computing [1] as new paradigm that emanates from

- Crowd wisdom / collective intelligence
- Crowdsourcing

Participatory sensing

- "It tasks average citizens and companioned mobile devices to form participatory sensor networks for local knowledge gathering and sharing."

and becomes "a new sensing paradigm that empowers ordinary citizens to contribute data sensed or generated from their mobile devices, aggregates and fuses the data in the cloud for crowd intelligence extraction and human-centric service delivery."

[1] Quoting from section 1, including Table 1, of [Guo et al., 2015]

Participation and awareness [Guo et al., 2015]

Three dimensions of awareness in mobile crowdsensing and computing [2]

- Sensing user context (location, activity) and behavioural patterns — User awareness
- Sensing the environment and its semantics — Ambient awareness
- Sensing the social group and its activities — Social awareness

where *user awareness* may be explicit or implicit, and *user involvement* may be participatory or opportunistic.

Mobile crowdsensing in the context of this tutorial:

Crowdsensing for medical / healthcare purposes with:

- active user involvement
- user as participating agent $\longrightarrow$ patient empowerment
- eventually with data sharing

[2] From section 2.4 and Fig. 1, change in notation

Personal sensing for mental health [Mohr et al., 2017]

Personal sensing for mental health: Goal

"...to convert the potentially large amount of raw sensor data into meaningful information related to behaviors, thoughts, emotions
(for simplicity, in this article we refer to these collectively as behavioral markers),
and clinical states and disorders."

Clinial state

"**Behavioral marker:** behaviors, thoughts, feelings, traits, or states identified using personal sensing"

"**Feature:** a measureable property of a phenomenon, which is proximal to, and constructed from, sensor data"

Raw sensor data

An example hierarchical framework [Mohr et al., 2017]

Figure removed

Clinical targets and personal sensing

Linking clinical targets to low-level features [Mohr et al., 2017] (Section 2.3)

- Behavioral markers: high-level features associated with a clinical state – defined by the medical expert
- Low-level features, whose association with a behavioral marker is built by combining ML/DM and human knowledge

Example [Mohr et al., 2017] (Sections 2.3, 3.1)

- Behavioral marker: Sleep disruption
- Low-level features: Bedtime/waketime, phone usage, movement intensity, ambient noise
- Clinical targets: depression, bipolar disorder, schizophrenia

From low-level-features towards clinical targets

Clinial targets: states, disorders – may be predicted or explained better with help of personal sensing

Caution on predicting clinical targets [Mohr et al., 2017] (Section 2.4)

"One would not attempt to diagnose a mental health disorder on the basis of one or two questions about symptoms (although one might use them for screening purposes). . . . We expect that clinical targets will be better predicted by applying machine- learning methods to a larger number of behavioral markers and features. "

Caution 2:

Personal sensing may contribute (a) to tracing the presence of the clinical state and (b) to understanding the clinical disorder.

These are two distinct tasks.

Challenge 1: Study quality and reproducibility

[Mohr et al., 2017] (Section 5.1)

Reproducibility of findings: is a must.

- Efforts to replicate the findings of earlier studies do exist BUT
- "…studies that appear to address the same behavioral marker usually use different sensors, different sets of features, different methods of measuring the behavioral markers, and varying research designs (e.g., giving people phones versus having them use their own, studying them for varying periods of time, or having varying numbers of participants excluded).

  The machine-learning methods used vary, and the results or weightings, particularly for group models, are not necessarily comparable across studies.

  In addition, it is unclear how many attempts have not been published due to failure."

↑ "because computer science and engineering tend to value technical novelty over generalizability"

Challenge 1 cntd [Mohr et al., 2017] (Section 5.1)

Quality of the results:
"...the availability of easy-to-use tools for machine learning is expanding faster than the expertise, resulting in a growing number of publications using questionable methods. ...
papers that used these inappropriate techniques were cited just as often as papers using proper techniques, suggesting that poor-quality information is having the same impact as high-quality information."

Challenge 2: Curse of Variability [Mohr et al., 2017] (Section 5.2)

- Sources of variability increase: characteristics of devices, characteristics of people, characteristics of social environments
- Pooling data across different studies together seems less easy than it sounds see also: reproducibility

On the positive side:

> *"The field of personal sensing in mental health is still young and small enough that some agreement*
> *on a core set of clinical assessment methods (EMA or self-report) may be possible,*
> *thereby providing uniform anchors to which the broad range of sensor data could be tied as it evolves and changes over time and across research projects."*

Challenge 3 on the Unknown Expiration Date

[Mohr et al., 2017] (Section 5.3)

- Technology becomes outdated.
- People change in their behavior.
- Technology contributes to changes in people's behavior.

cf. Google Flu Trends [which] "mined flu-related search terms . . . It performed remarkably well until it stopped working. . . . The changes in people's search strategies were driven at least in part by Google's own efforts to optimize search algorithms, which also altered the search recommendations provided to users, thus changing people's search behaviors . . . "

Challenge 4 on Accuracy vs Invisibility [Mohr et al., 2017] (Section 5.4)

Personal sensing makes unobtrusive data acquisition possible.

- Requiring user action is often a burden and makes people give up the tool
- User inputs deliver context and can make the tool more personalized and more accurate.

Making the ends meet:

> *"Rather than thinking of a sensing platform as a technology that autonomously creates information, it may be more useful to think of the sensing platform as a social machine in which the quality of prediction reflects a shared endeavor."*

Challenge 6 on Privacy and Trust [Mohr et al., 2017] (Section 5.6)

- Privacy, and the right of people over their own data
- De-anonymization, an easy task (relatively few data needed, esp if GPS traces are among them)
- Privacy management and trusting it

Challenge 5 of Uncertainty [Mohr et al., 2017] (Section 5.5)

- Sensitivity and Specificity
- False alarms and false negatives

in the context of personal sensing – for a specific patient and clinical staff

WHERE to use personal sensing?

- [a] Integration in existing models of care ⇒
  - ! establishing workflows and infrastructures that link personal sensing in existing models of care
  - ! making sure that clinical staff and caretakers know how to respond to alerts raised by a personal sensing system
- Behavioral interventions – cf. studies on how CBT can be delivered with help of internet and e-technologies
- Epidemiology – using personal sensing to understand clinical disorders

[a] Black texts come from [Mohr et al., 2017], Section 6, but gray texts do not.

Analyzing mHealth data to understand tinnitus

[Probst et al., 2017a]

Why monitor tinnitus?

- 5.1% to 42.7% of the population experience tinnitus (citing McCormack et al (2016)).
- Some tinnitus patients experience stress, depression, anxiety, fatigue, insomnia, some become even incapable of working.
- Cognitive Behavioural Therapy (CBT) has been shown to reduce the burden of tinnitus
- but patient response to treatment varies – to CBT and, even more, to other forms of treatment.
- One explanation for poor response and inconsistent results is *heterogeneity*:
  - inter-individual heterogeneity: tinnitus varies across patients
  - inter-individual heterogeneity: tinnitus for a patient varies over time

Remembering tinnitus symptoms [Pryss et al., 2017]

Remembering – why?

"The treatment of tinnitus and the early diagnosis of potential comorbidities require assessments on several symptoms, including loudness and variation of the perceived sound(s), distress caused by tinnitus, impact of tinnitus on sleeping behavior, comorbidities, social activity, concentration, and so forth."

Remembering – how well?

"Bratland-Sanda et al (2010) [5] assessed physical activities of patients with eating disorders by retrospective self-reports as well as . . . with an accelerometer. Patients reported significantly less physical activity retrospectively than what was measured prospectively by the accelerometer."

Recording instead of remembering

Ecological Momentary Assessments (EMA): observable variables (e.g. symptoms) are repeatedly *assessed* – citing Trull & Ebner-Priemer (2013) on "ambulatory assessments"

EMA using smartphones [Probst et al., 2017a, Pryss et al., 2017]

TrackYourTinnitus mobile app:

Registration for tinnitus monitoring

Three questionnaires:

- Mini-TQ-12 on tinnitus-related psychological problems
- TSCHQ (37) on tinnitus sample case history
- Worst Symptom Questionnaire (9)

to be filled once.

EMA on tinnitus

Seven questions on:

- tinnitus loudness
- distress through tinnitus
- valence and arousal

to be answered up to 12 times a day at randomly chosen moments.

- Ambient sounds are captured during each EMA recording.

Analyzing EMA on TrackYourTinnitus [Probst et al., 2017a]

Time-of-day dependence of tinnitus loudness and distress:

Materials

- Total assessments: 25,863 Retained: 17,209, after excluding assessments with missing values in any of the target variables and days with less than three assessments.
- 350 participants (253m/94f) with average age 45.4 (over 333, SD=12.1) and median since tinnitus onset 5.4Y (from 0 to 61.8Y)
- Median days per participant 11 (from 1 to 415) with median number of assessments per day 4 (from 3 to 18)

Specifying day and night intervals

- night: 12am–4am
- early morning: 4am–8am
- early evening: 4pm–8pm

afternoon: 12pm-4pm
late morning: 8am–12pm
late evening: 8pm–12am

Analyzing EMA on TrackYourTinnitus [Probst et al., 2017a]

Time-of-day dependence of tinnitus loudness and distress:
Selection of findings

- “tinnitus was significantly louder in the late evening compared to the afternoon and early evening.”
- “stress-level increased from morning to afternoon, decreased from afternoon to evening, and did not differ compared to the night”
- “Tinnitus was louder and more distressing when the level of stress was higher at a specific time-of- day compared to other times-of-day, when it was higher during a whole day compared to other days, and when it was higher during the whole assessment period for a given participant (compared to other participants).”
- “the effects of time-of-day on tinnitus loudness and tinnitus distress were still significant (i.e., after controlling for the effects of stress).”

Reaching the patients

[Probst et al., 2017b]

Same disease, same population?

Three sources

Source	*Type*	*Sample size*
Tinnitus Center of Univ Hospital Regensburg	outpatient clinic	3786
TrackYourTinnitus	mobile app	867
TinnitusTalk	self-help social platform	5017

Results of comparison

Significant differences in age, gender and time since tinnitus onset $(p < 0.05)$

- Age: TrackYourTinnitus users were younger
- Gender: more female users in TinnitusTalk
- Time since tinnitus onset: users of TrackYourTinnitus & TinnitusTalk had more often acute, resp. subacute tinnitus (less than 3M, resp. 4-6M) or tinnitus for more than 20Y

Understanding the patients: different media, different needs

TrackYourTinnitus: Clustering patient evolution [Unnikrishnan, 2017]

Figure removed

Monitoring opinions on treatments in TinnitusTalk [Dandage et al., 2017]

Figure removed

Tinnitus Center Univ Hospital Regensburg [Schneck et al., 2017]

Finding questionnaire entries that capture the loudness/handicap interplay

		Top-10 variables for L_H+
Both	8	THI:{Q10,Q12,Q13,Q16,Q17}, TQ:{Q7,Q10,Q15}
MT_{RF}	2	THI:{Q1,Q23}
LP_{RF}	2	THI:Q21,TQ:Q39

		Top-10 variables for $L+H\_$
Both	7	THI:{Q10,Q12,Q13,Q16,Q17}, TQ:{Q7,Q15}
MT_{RF}	3	THI:{Q1,Q15,Q25}
LP_{RF}	3	THI:{Q7,Q14,Q21}

Learning from Mobile Data - Outline

- $\surd$ Mobile sensing and crowdsensing
- $\surd$ Mobile sensing for mental health
- $\surd$ Mobile sensing and Ecological Momentary Assessments

Achievements and Open Challenges

- The challenge of finding the data
- The challenge of seeing with the expert's eyes
- The challenge of preparing the data
- Challenges of learning
- The challenge of explaining the results

THANK YOU !

KMD LAB

- Classification, clustering, semi-supervised and active machine learning
- Maching learning on texts: analysing sentiment, predicting polarity
- Stream mining and model adaption to drift
- Adaptive recommendation engines

. . . dealing with big data

- ⋆ **Volatility:** adaptive learning methods
- ⋆ **Veracity:** semi-supervised/constraint-based/active learning
- ⋆ **Sparsity:** learning on data with gaps

. . . and the same on small data

VISIT THE KMD LAB:

- http://www.kmd.ovgu.de/
- Faculty of Computer Science, Otto-von-Guericke-University Magdeburg
- Sendmail at: myra@ovgu.de

- Thank you!

Acknowledgements: German Research Foundation project OSCAR
"Opinion Stream Classification with Ensembles and Active Learners"

Bibliography

[Dandage et al., 2017] Dandage, S., Huber, J., Janki, A., Niemann, U., Pryss, R., Reichert, M., Harrison, S., Vessala, M., Schlee, W., Probst, T., and Spiliopoulou, M. (2017).
Patient empowerment through summarization of discussion threads on treatments in a patient self-help forum.
In *Proc. of Int. Conf. on Biomedical and Health Informatics (ICBHI 2017)*, Thessaloniki, Greece.

[Glenn, 2005] Glenn, N. D. (2005).
Cohort Analysis.
Quantitative Applications in the Social Sciences. SAGE, 2nd edition.

[Guo et al., 2015] Guo, B., Wang, Z., Yu, Z., Wang, Y., Yen, N. Y., Huang, R., and Zhou, X. (2015).
Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm.
ACM Computing Surveys (CSUR), 48(1):7.

[Mohr et al., 2017] Mohr, D. C., Zhang, M., and Schueller, S. M. (2017).
Personal sensing: understanding mental health using ubiquitous sensors and machine learning.
Annual review of clinical psychology, 13:23–47.

[Probst et al., 2017a] Probst, T., Pryss, R. C., Langguth, B., Rauschecker, J. P., Schobel, J., Reichert, M., Spiliopoulou, M., Schlee, W., and Zimmermann, J. (2017a).
Does tinnitus depend on time-of-day? an ecological momentary assessment study with the “trackyourtinnitus” application.
Frontiers in Aging Neuroscience, 9:253.

Bibliography

[Probst et al., 2017b] Probst, T., Pryss, R. C., Langguth, B., Spiliopoulou, M., Landgrebe, M., Vesala, M., Harrison, S., Schobel, J., Reichert, M., Stach, M., and Schlee, W. (2017b).
Outpatient tinnitus clinic, self-help web platform, or mobile application to recruit tinnitus study samples?
Frontiers in Aging Neuroscience, 9:113.

[Pryss et al., 2017] Pryss, R. C., Probst, T., Schlee, W., Schobel, J., Langguth, B., Neff, P., Spiliopoulou, M., and Reichert, M. (2017).
Mobile crowdsensing for the juxtaposition of realtime assessments and retrospective reporting for neuropsychiatric symptoms.
In *Proc. of IEEE Symposium on Computer-Based Medical Systems (CBMS 2017)*, Thessaloniki, Greece.

[Schneck et al., 2017] Schneck, A., Kalle, S., Pryss, R., Schlee, W., Probst, T., Langguth, B., Landgrebe, M., Reichert, M., and Spiliopoulou, M. (2017).
Studying the potential of multi-target classification to characterize combinations of classes with skewed distribution.
In *Proc. of IEEE Symposium on Computer-Based Medical Systems*, Thessaloniki, Greece.

[Unnikrishnan, 2017] Unnikrishnan, V. (2017).
Analysis of patient evolution on time series of different lengths.
Faculty of Computer Science, Otto-von-Guericke Univ. Magdeburg.
Master Thesis.