

Diagnostic paths - towards computational evidence

Liniger Z¹, Lindner G², Fiedler GM¹, Leichtle AB¹

[1] Center of Laboratory Medicine - University Institute of Clinical Chemistry and [2] Department of Emergency Medicine at the Inselspital – Bern University Hospital, Switzerland

Abstract

Background

For the establishment of rapid, efficient, and financeable laboratory diagnostics in an ICD-10-funded hospital environment, defined diagnostic paths are of ever-growing importance. Usually they are set up as hierarchical trees or flow charts, leading to altered diagnostic suggestions depending on the outcome of previous tests (step-by-step diagnostic schemes). These guidelines are designed to focus diagnostic efforts on the most selective analytes, thereby avoiding unnecessary testing and providing a test panel sufficient to cover the most important side diagnoses. By overall reducing the number of recommended tests, they can help to improve cost effectiveness - especially in a health care compensation system that includes all diagnostic testing in a flat charge (as e.g. the recently introduced Swiss DRG system).

Diagnostic paths arise from different sources: single publications, recommendations of the different medical associations, and from the (inter-)national societies of Laboratory Medicine. However, all of these recommendations bear a severe drawback: they are agreements of experts in the field and therefore reflect opinions, not evidence.

Methods

Along with the rapid evolution of computational tools for parallel, GPU-based and grid- or cloud-based computing and the development of powerful statistical strategies an appealing resolution for this aforementioned lack of evidence emerges: To utilize already collected laboratory data, to merge it with diagnostic information and to derive which analyses are selective and likewise superior for the setup of a diagnosis. The conventional diagnostic path recommendations can be replaced by diagnosis-specific models inferred from the laboratory and classification data, a process that does not imply prior "expert knowledge" in terms of opinions which parameter to measure, but that is based only of the numerical evidence contained in the dataset and that delivers disease probabilities, assisting the physician to balance decisions for further diagnostic and therapeutic procedures. In a proof-of-principle study we utilized laboratory data and diagnostic information of n=15'000 patients of the Inselspital's Department of Emergency Medicine and computationally determined, which lab tests are indispensable and which ones are unnecessary for establishing the top ICD-10 coded diagnoses via the estimation of posterior inclusion probabilities with bootstrapped confidence intervals of a set of lab tests.

Results

Our results clearly show the feasibility of our new approach: For myocardial infarction e.g. our algorithm without any prior knowledge of the disease nor any pathophysiological basis suggests a panel of lab test similar to current guidelines - solely based on computational principles, our patient population, and laboratory data already generated thereof.

Conclusion

In a highly digitalized hospital environment, the present lack of evidence for diagnostic paths is unjustifiable: All diagnosis-related classification of all patients - hospitalized or out-patient- are electronically registered, and usually all lab tests are also electronically available. These data, stored away and laid untouched for decades, could improve and streamline diagnostic testing and implicitly generate benefit for the patients - the tools therefor are ready.

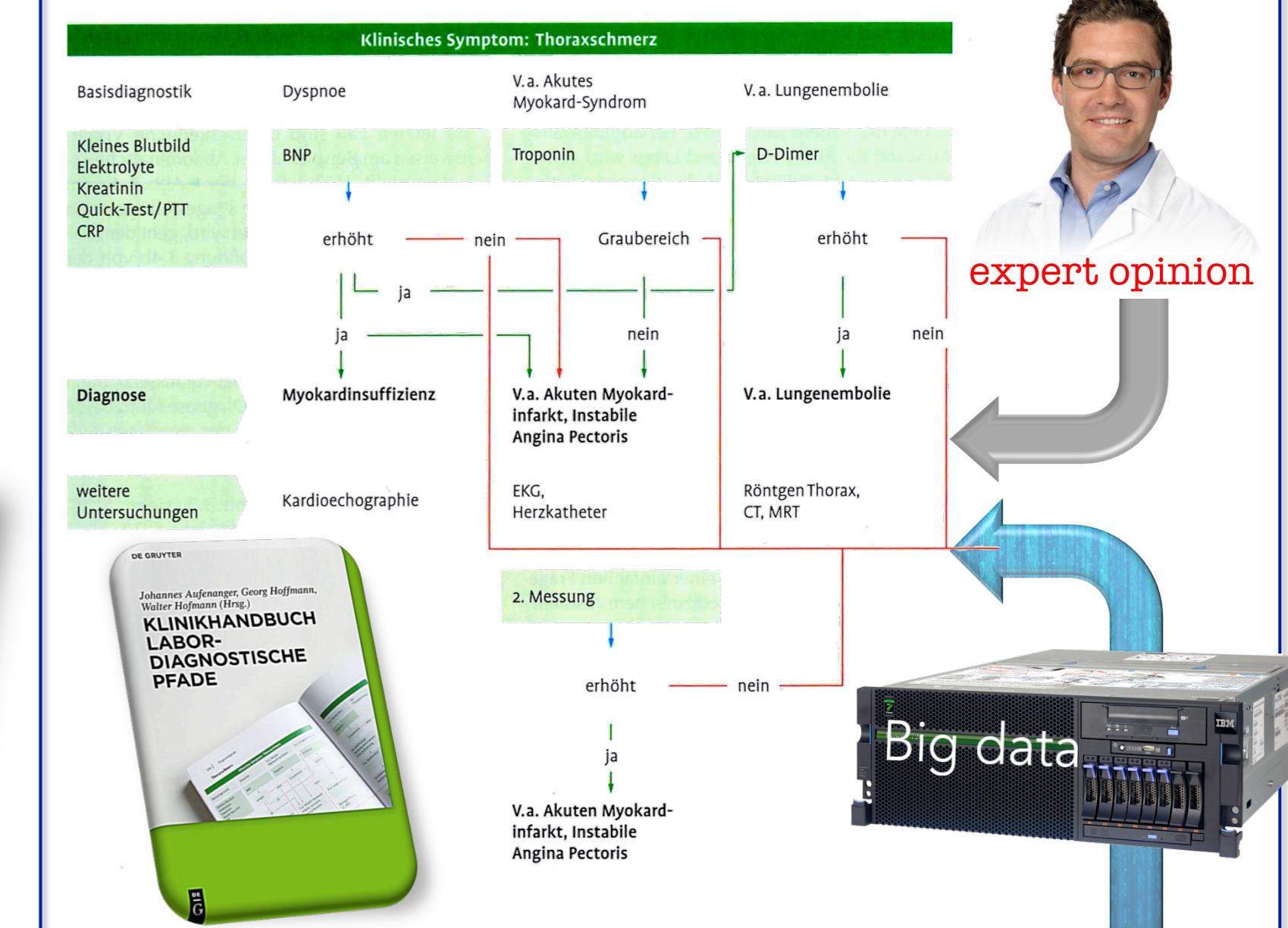


Introduction

Selecting the right tests for differential diagnosis – a matter of chance?
 Today's lab request systems offer only limited guidance for selecting adequate diagnostic procedures depending on potential side diagnoses and with respect to clinical effectiveness and economic reasonability.
 Therefore, the actual selection of lab tests depends more on the experience, specialization, and textbook adherence of the requesting physicians than on evidence and rationality.



Objective



Our Objective: Replacing „expert opinion“ by computational evidence
 Currently, clinical diagnostic guidelines are mainly based on recommendations and agreements of expert boards and thought leaders. In times of “big data” this lack of scientific evidence cannot be justified anymore. Everything needed is at hand: the data, the algorithms, and the computational power.

Figures/Images: Aufenanger et al., De Gruyter, 2012 and IBM™

Prerequisites



Diagnoses and laboratory data



Computational power



Algorithms and time

The prerequisites for the computational generation of diagnostic recommendations

Since the guidelines should be based on evidence from data and not on opinions, we generated a data base of all admission lab test results of the patients of the Inselspital's emergency ward from 2009 and 2010, limited on the 20 most frequent test requests (n>15'000), and combined them with the ICD-10 coded main diagnoses.

For the computations, we use a Xeon® E5-2620 based workstation equipped with 64G RAM running Scientific Linux 6.5. Parts of the imputation were optimized in collaboration with Peter Weinert from the German Supercomputing Center at the Leibniz Rechenzentrum in Garching, Germany. We also run part of our R scripts on the UBELEXI linux cluster of the University of Bern, Switzerland.

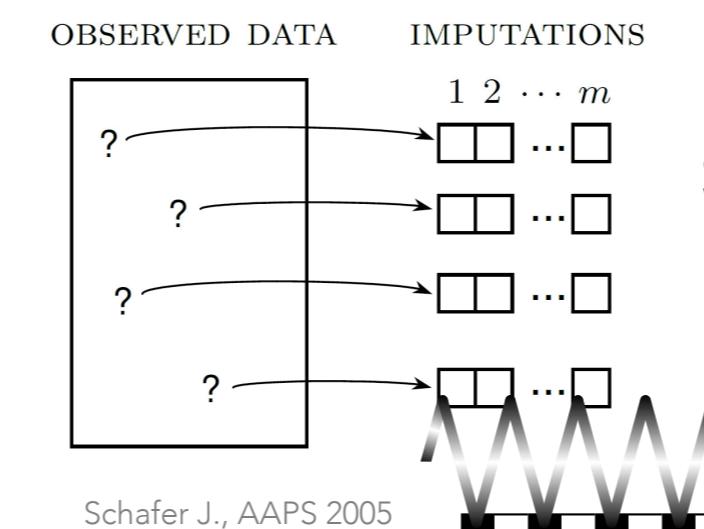
Our computational pipeline consists of a series of consecutive R scripts (R v. 3.1.0) invoking also C scripts for the evaluation of model probabilities. Especially the imputation steps are extremely resource intensive: our workstation needed 5101303 seconds (~59d) for 2000 imputation cycles on a restricted dataset (n=3425).



Imputation

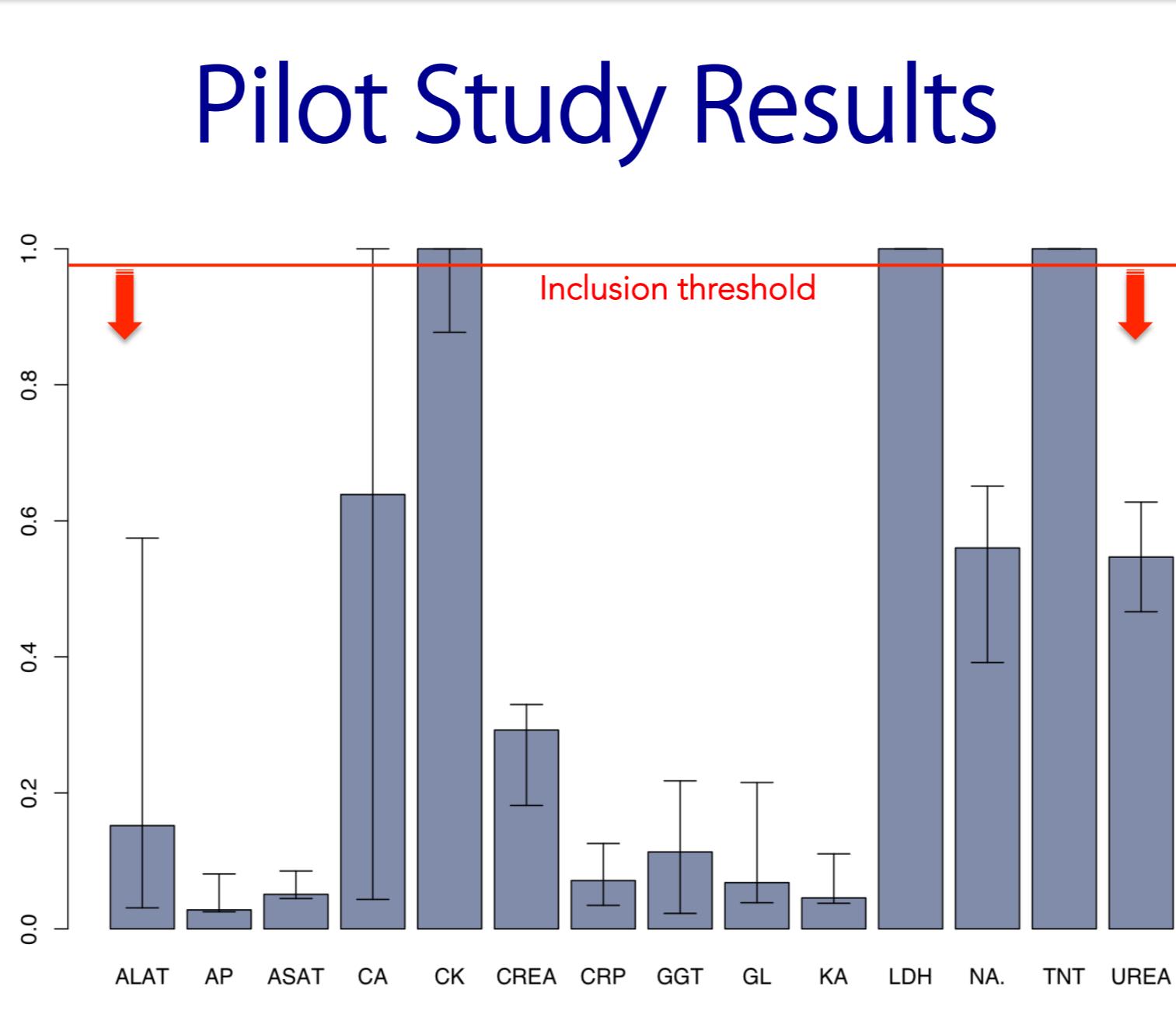
MDIA	KA	NA.	TNT	UREA	LDH	ALAT
A084	3.1	139	0.009	6.6	363	13
A084	4.1	132	0.01	15.1	1023	41
A084	3.2	136	0.009	20.7	NA	25
A09	3.7	135	0.009	6.8	NA	17
A09	2.9	138	0.012	7.2	NA	14
A09	4.5	147	0.009	17	505	37
A09	3.5	139	0.009	18.1	NA	NA

Multiple Imputation
 We chose the well-established multiple imputation procedure (MI) to replace missing data - a MCMC-simulation based approach also referenced as repeated-imputation inference method. MI imputes each missing value with several separate imputed results ("chains"). It starts with a crude "random" imputation and loops through each "iteration" while it takes the matrix of observed variables as regression predictors X and the prior imputed matrix Y of missing data as basis for the next imputation step - until approximate convergence.



Missing data
 As we acquire "real" clinical lab results as data basis, the request profiles are inconsistent, leading to a considerable amount of "missing data" (NAs). To make use of as many cases as possible, we limited our data set on the "Top 20" requests and imputed the missing results.

MDIA	KA	NA.	TNT	UREA	LDH	ALAT
A084	3.1	139	0.009	6.6	363	13
A084	4.1	132	0.01	15.1	1023	41
A084	3.2	136	0.009	20.7	281.63	25
A09	3.7	135	0.009	6.8	332.72	17
A09	2.9	138	0.012	7.2	411.44	14
A09	4.5	147	0.009	17	505	37
A09	3.5	139	0.009	18.1	528.06	50.85



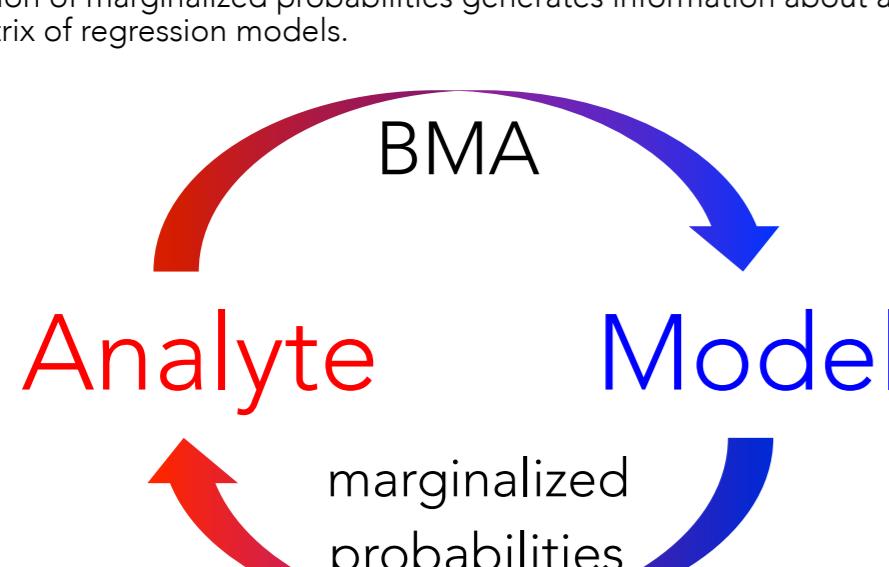
Results of a pilot study on Myocardial Ischemia parallel conventional expert recommendations
 In a small pilot study (subset of the complete dataset, consisting of 459 cases of myocardial ischemia (ICD-10: I200-I259) and 2966 controls) we determined the analyte inclusion probabilities of 14 analytes with their respective confidence intervals. The results impressingly show that without prior knowledge the algorithms identify the same lab tests as "valuable" as conventional clinical recommendations do. For establishing stepwise diagnosis schemes, a sliding "inclusion threshold" could be applied at the diagnosing physicians' discretion or budget.

Legend:
 ALAT alanine transaminase; AP alkaline phosphatase; ASAT aspartate transaminase; CA calcium; CK creatin kinase; CREA creatinine (enz); CRP C-reactive protein; GGT gamma-glutamyl transpeptidase; GL glucose; KA potassium; LDH lactate dehydrogenase; NA sodium; TNT troponin T; UREA urea.

The y-axis shows inclusion probabilities (0 ≈ 0%, 1 = 100%).

A Recursive Approach

The computational evaluation of an analyte's diagnostic value
 How do we proceed from diagnostic models containing different sets of variables to an analyte-based evaluation of diagnostic value?
 The basic assumption is: If an analyte is included in many models, its diagnostic value is probably high. Via the computation of (marginalized) model inclusion probabilities of each analyte out of the model matrix (see the plot at the right edge of this box, displaying analytes in columns and models in rows as well as inclusion in dark and non-inclusion in bright), we get estimates of each analyte's probable diagnostic value, independent from the different models and variable sets these estimates were computed from. This approach is the next consequential, but likewise recursive step: Whereas BMA generates models of diagnostic values out of sets of analytes, the generation of marginalized probabilities generates information about analyte value out of a matrix of regression models.



The generation of confidence
 In this analytical pipeline uncertainty is omnipresent. It's implicated in the measurement of the analytes, in the imputation, in the modeling and in the estimation of the analytes' marginal probabilities. To estimate at least part of this uncertainty, we computed confidence intervals based on separate evaluations of the different chains of imputation (including modeling and probability estimation) by bias-corrected accelerated bootstrapping. The less observed "real" test results are e.g. observed (resulting in more imputed values), the larger the confidence intervals might expand. Nevertheless, also the certainty of model inclusion contributes substantially to the confidence intervals (e.g. the difference between GGT and LDH with comparable request frequency).

Pilot Study Results

CONCLUSIONS

„Big data“ offers the chance to replace conventional diagnostic recommendations by a computationally evidenced guidance system.

The encouraging results of our proof-of-concept study demonstrate the principal feasibility of this approach, which is applicable to a vast variety of diagnostic questions.

Collaborations



Ethics & Disclosures
 According to the Swiss law on human research (Law Nr. 810.30 Art. 2 [2c]), this study is based on routinely generated, completely anonymized data and is therefore not subject to approval by the Cantonal Ethics Committee (dispensation granted Nr. 2014/2023). No relevant disclosures exist.

