

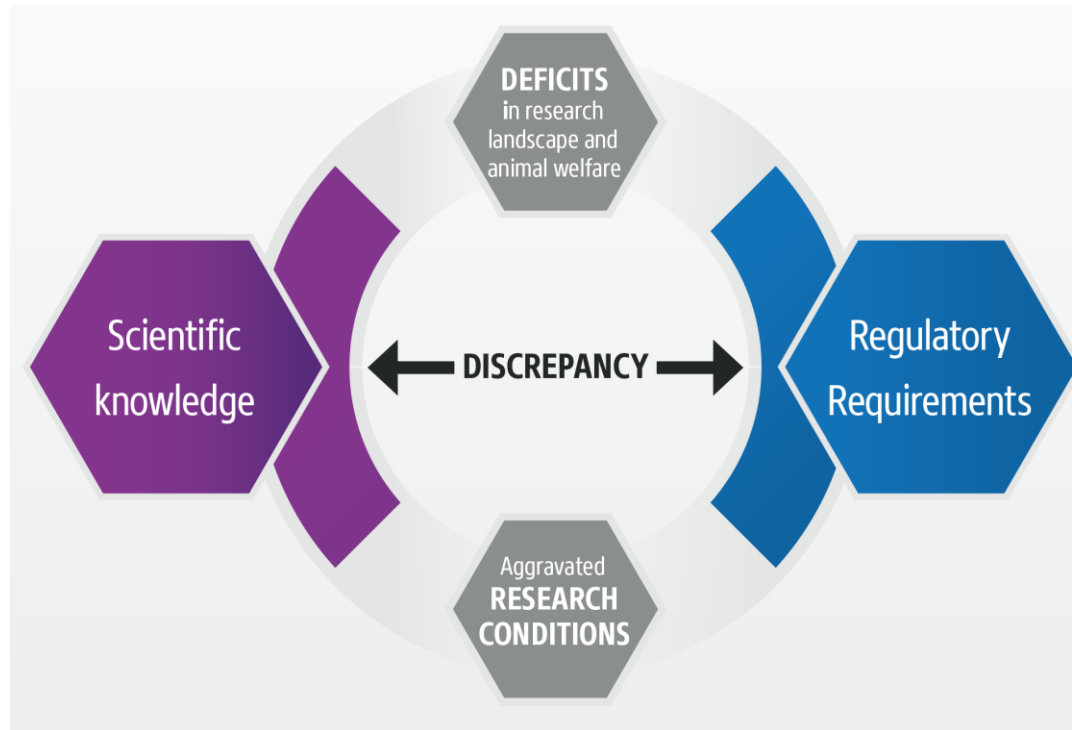


Severity assessment and refinement in bile duct ligation models

Dietmar Zechner
Rudolf-Zenker-Institute
for Experimental Surgery
Rostock University Medical Center

FOR 2591: Severity Assessment
in animal based research

Status quo



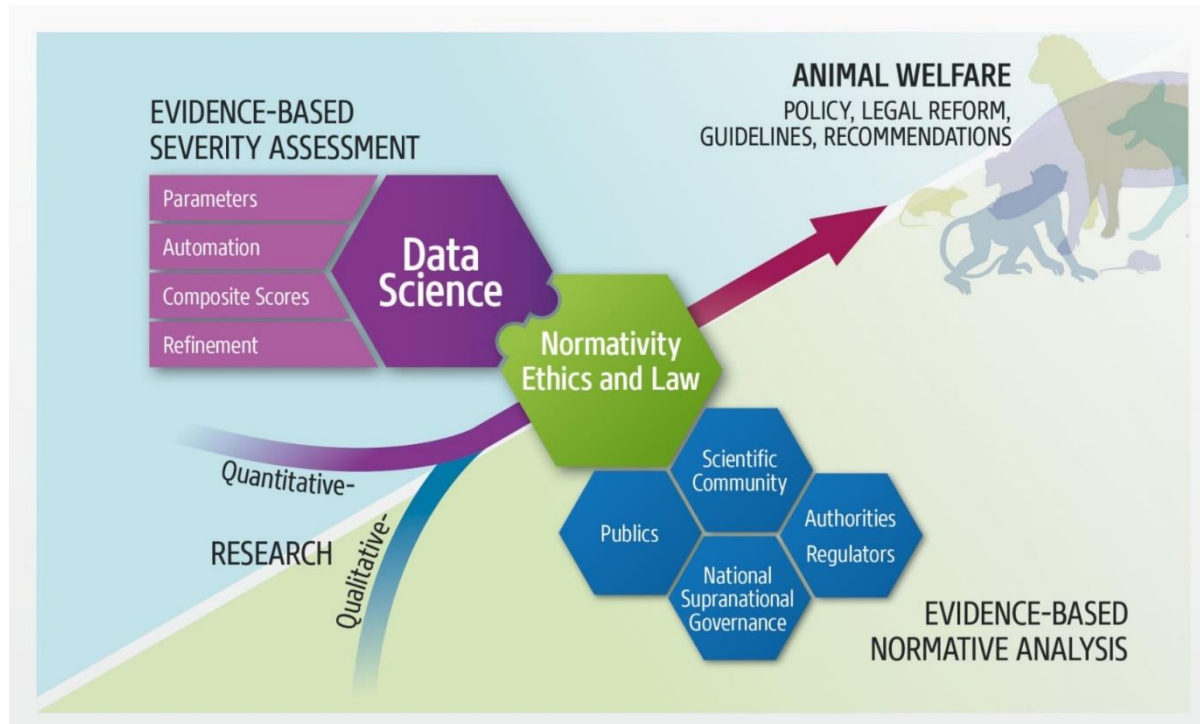
Science

- which methods are useful to assess severity?
- which mathematical methods are appropriate?
- can we compare severity between animal models?
- how do we measure refinement?

Regulatory requirements

- prospective severity assessment
- retrospective severity assessment
- continuous refinement

Goal



Improve animal wellbeing based on

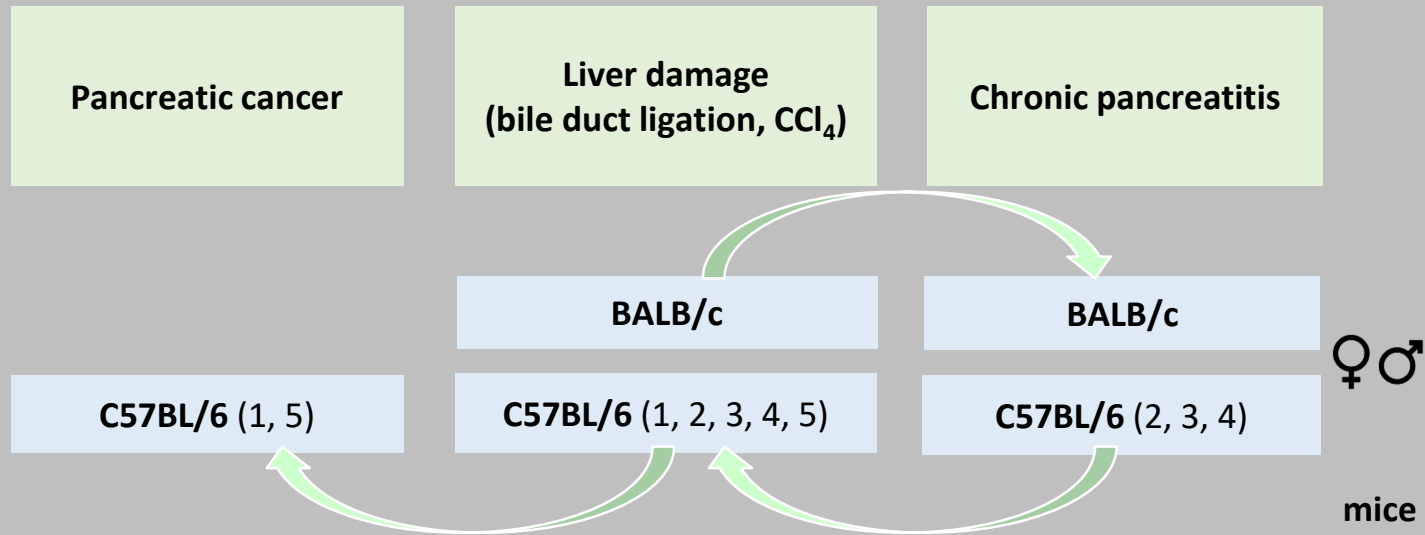
- interaction: Public - Scientific Community - Authorities
- data

Projects pursued in Rostock

FOR 2591
Zechner
Vollmar

DFG FOR 2591

Evaluate & compare distress in mice (drug versus vehicle control)



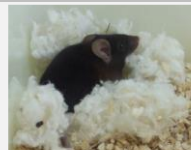
Refinement

- 1: Optimise humane endpoints (Cox proportional hazards model)
- 2: Reduce pain by evaluating analgesics
- 3: Optimise different surgical interventions

Novel methods

- 3: Thermal imaging (inflammation, heart & respiratory rate,...)
- 4: Automated video analysis (running activity, grimace scale,...)
- 5: Volatile components as biomarkers

Main read out parameters



+

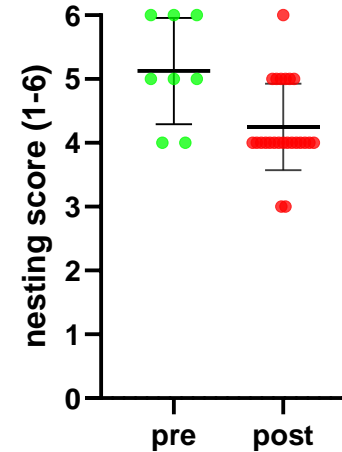
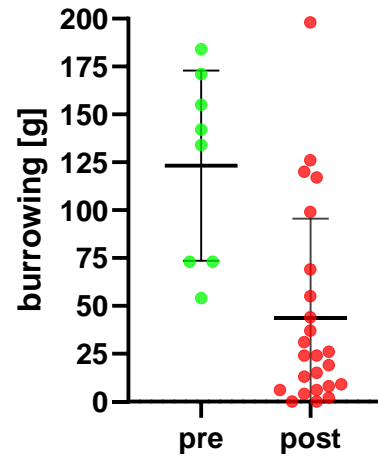


Which methods can measure distress?

Indirect interpretations

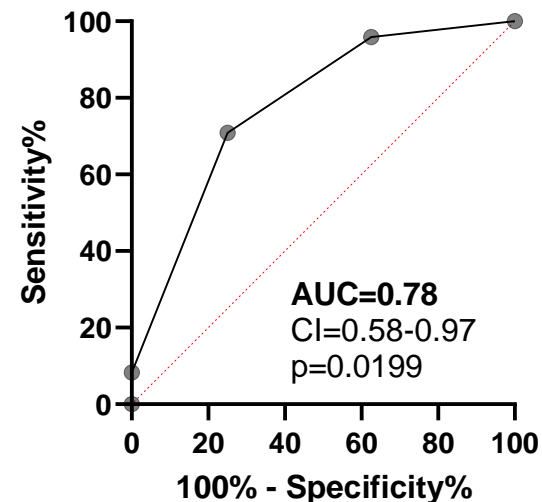
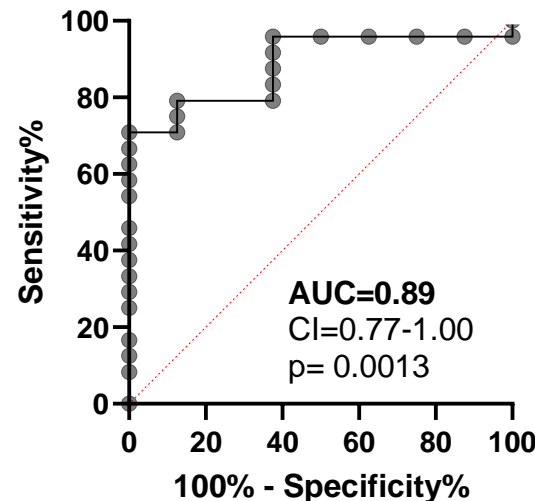
(limitation:
misinterpretation?)

- healthy vs. diseased
- receiver operating characteristic (ROC) curves

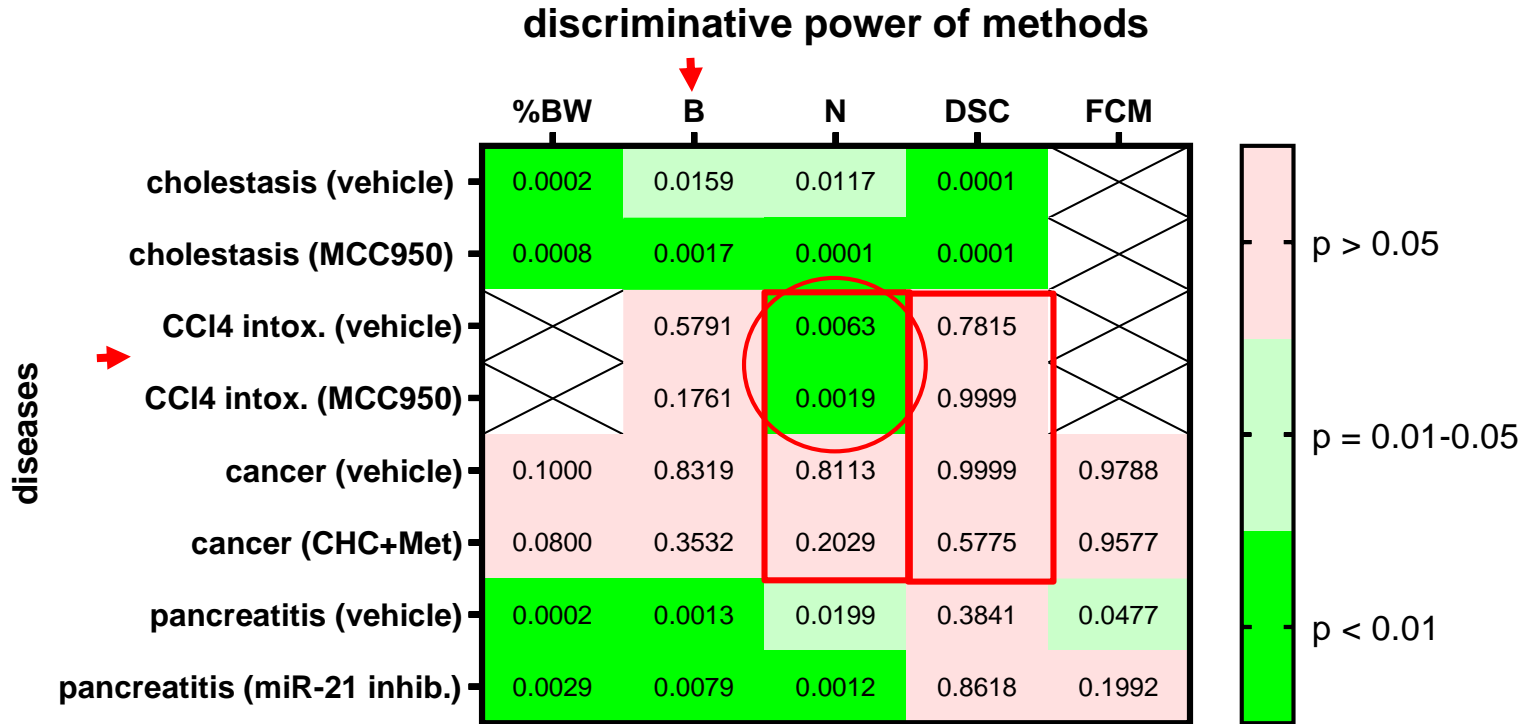


chronic pancreatitis

- 50 µg/kg, cerulein i.p.
- 3x/day; 3x/week
- for 4 weeks
- analyse: pre, early, middle and late phase



Which methods can measure distress?



Conclusions

- Different methods can measure differences in different animal models
- the **methods** seem to be **fairly robust** (within an animal model the same methods can often differentiate between healthy and diseased animals independent if animals are treated with a drug or not)
- Use more than 1 method! → **Multivariate analysis**

How to compare animal models?

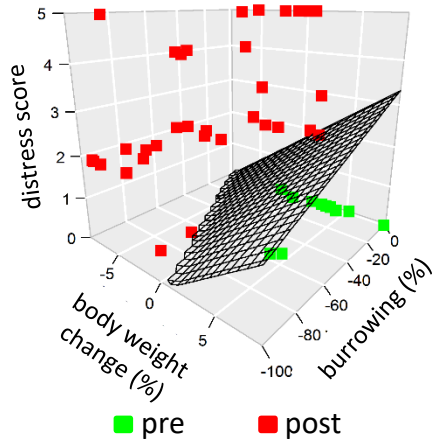
Training data set: laparotomy (pre vs. post, moderate distress?)

Methods combined: % body weight, burrowing and distress score

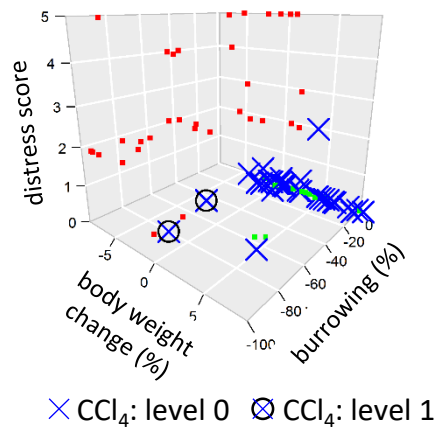
Multivariate analysis: Support Vector Machine classification

Compare: during CCl₄ induced liver damage vs. after bile duct ligation (BDL)
(supervised machine learning)

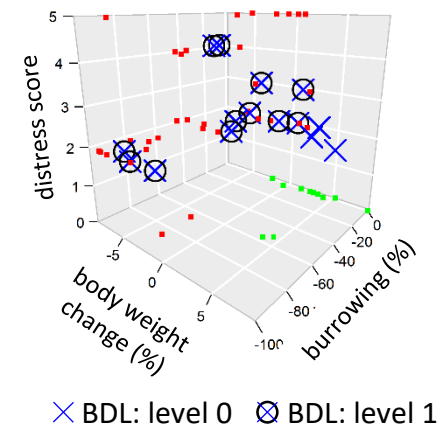
Training data: Pre and post-laparotomy



Test data: CCl₄



Test data: BDL



Body weight & Burrowing & distress score	CCl ₄	BDL
distress level 0 (pre laparotomy)	28	3
distress level 1 (post laparotomy)	2	12
Significantly different distribution of distress levels between CCl ₄ and BDL (Fisher's Exact Test, $P \leq 0.001$)		

Conclusion

- CCl₄ ≤ BDL

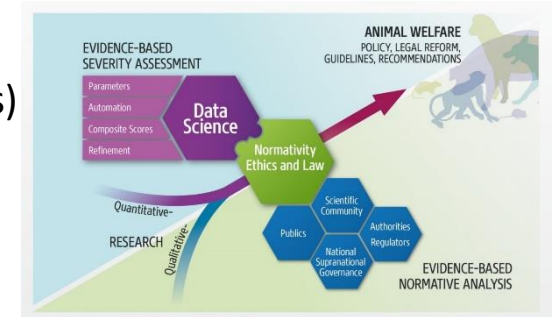
Conclusion/Discussion

What can science do for animal welfare?

- **Compare distress** between animal models (use multivariate analysis)
- **Refine animal models** (use multivariate analysis)

limitations:

Do methods really always measure animal welfare?



Open questions

- **Replicability** and **robustness** of distress evaluation?
 - can person A replicate the data of person B?
 - does an animal model always have the same severity? (Δ mouse strains, age, SPF status)
 - How do variations of animal models influence distress?
 - How do different analgesics influence distress?
- Is it scientifically sound to assume **3 levels of severity**?
 - Or will all animal models represent a continuum of severity levels and any border between severity levels will be arbitrary? Or will severity levels cluster in more than three categories?
- How exactly should we value **cumulative suffering**?

Animal model X
≠
always
severity category Y