



Weierstrass Institute for  
Applied Analysis and Stochastics

# Workshop on Validation in Statistics and Machine Learning

Nicole Krämer   Anne-Laure Boulesteix



- **Weierstrass Institute for Applied Analysis and Stochastics**
  - part of the Leibniz association
- 
- independent, publicly funded research institute
  - $\approx$  120 members of staff
  - 7 research groups, 2 Leibniz groups
  - object-oriented research in applied analysis and stochastics



**Weierstrass Institute for  
Applied Analysis and Stochastics**



**PASCAL2**

Pattern Analysis, Statistical Modelling and  
Computational Learning

*How do we convince ourselves and others that our empirical findings are valid?*

1. **no ground truth** in many real-world applications
2. difficulties to **reproduce** published results
3. **peer pressure** to only publish positive findings and significant improvements
4. pre-publication optimization of the algorithms and **publication bias**
5. high **instability** of empirical results
6. ...

*How do we convince ourselves and others that our empirical findings are valid?*

1. **no ground truth** in many real-world applications
2. difficulties to **reproduce** published results
3. **peer pressure** to only publish positive findings and significant improvements
4. pre-publication optimization of the algorithms and **publication bias**
5. high **instability** of empirical results
6. ...

- 13 invited talks
- 3 contributed talks
- 8 poster presentations **(this afternoon!)**
- > 50 participants
- lively discussions after the talks, in the coffee and lunch breaks, and at the conference dinner

Focus on biological and medical applications, but:

Validation is highly important in **all** applications  
of statistics and machine learning!!

Wednesday, October 6, 2010	
09.00 – 09.15	Welcome
<b>Session: “Telling the Story”</b>	
10.45 – 11.15	Coffee break
<b>Session: “ Variable Importance”</b>	
12.20 – 14.00	Lunch Break
<b>Session: “Reproducibility ”</b>	
15.30 – 17.00	Poster Session and Coffee Break
<b>Session: “Assessing Prediction Models”</b>	
20.00 – 24.00	Conference Dinner at “Tapas y Más Tres”