

ICATPP 2011  
Como, 3-7 October 2011

# Quantifying the unknown

Issues in simulation validation  
and their experimental impact

Maria Grazia Pia

*INFN Genova, Italy*

Matej Batic, Marcia Begalli, Mincheol Han, Steffen Hauf, Gabriela Hoff, Chan Hyeung Kim,  
Markus Kuster, Lina Quintieri, Paolo Saracco, Hee Seo, Georg Weidenspointner, Andreas Zoglauer

*State University Rio de Janeiro, Brazil*

*Tech. Univ. Darmstadt, Germany*

*MPI Halbleiterlabor, Munich, Germany*

*XFEL, Hamburg, Germany*

*INFN Sezione di Genova, Italy*

*INFN Laboratori Nazionali di Frascati, Italy*

*Hanyang University, Seoul, Korea*

*UC Berkeley, USA*



**Shall I trust my simulation?**

**How much can I trust my simulation?**

**Does it have predictive value?**

**EGS5**

**MCNPX**

**Tripoli**

**PHITS**

**EGSnrc**

**Validation**

**MARS**

**Penelope**

**MCNP**

**SHIELD-HIT**

**FLUKA**

**Geant4**

**Intrinsic limits to simulation validation**

# Epistemic uncertainties

Epistemic uncertainties originate from **lack of knowledge**

Relatively scarce attention so far in Monte Carlo simulation  
Studies in deterministic simulation (*especially for critical applications*)

## **Possible sources** in Monte Carlo simulation

- incomplete understanding of fundamental physics processes, or practical inability to treat them thoroughly
- non-existent or conflicting experimental data for a physical parameter or model (for validation)
- applying a physics model beyond the experimental conditions in which its validity has been demonstrated

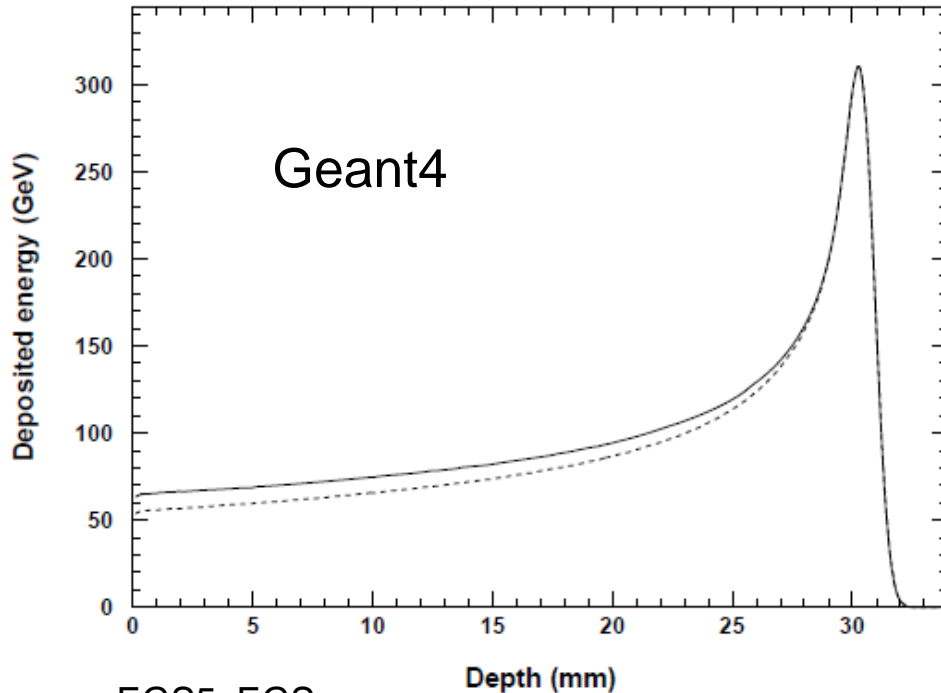
Epistemic uncertainties affect the **reliability** of simulation results

**Can we quantify them?**

# Uncertainty quantification

- Epistemic uncertainties are **difficult to quantify**
  - due to their intrinsic nature
- No generally accepted method of measuring epistemic uncertainties
  - and their contributions to reliability estimation
- Various formalisms developed in the field of deterministic simulation
  - **Interval analysis**
  - **Dempster-Shafer theory of evidence**
- Not always directly applicable in Monte Carlo simulation
  - Adapt, reinterpret, reformulate existing formalisms
  - Develop new ones specific to Monte Carlo simulation

# Warm-up exercise



- EGS5, EGSnrc
- Penelope
- MCNP(X)
- PHITS
- SHIELD-HIT
- FLUKA
- SPAR, CALOR, CEM, LAHET, INUCL, GHEISHA, Liège INCL, Bertini

- $p$  stopping powers
- Water ionisation potential
- $\delta$ -ray production
- Multiple scattering
- Nuclear elastic
- Nuclear inelastic
  - Cross sections
  - Preequilibrium
  - Nuclear deexcitation
  - Intranuclear cascade

$\delta$ -ray or **no  $\delta$ -ray**  
Preequilibrium or **no preequilibrium**  
Weisskopf-Ewing or **Weisskopf-Ewing**  
Griffin-exciton or **hybrid**  
etc.

# “Validation” in the literature

- Beam energy (and energy spread) is not usually known with adequate precision in therapeutical beam lines
  - What matters in clinical applications is the range
- Typical procedure: optimize the beam parameters to be used in the simulation by fitting them to experimental data
  - Determine beam energy, energy spread etc.
  - Use optimized beam parameter values in the simulation
- This is a **calibration**
- This is **NOT** validation

T. G. Trucano, L. P. Swiler, T. Igusa, W. L. Oberkampf, and M. Pilch,  
“Calibration, validation, and sensitivity analysis: What’s what”,  
*Reliab. Eng. Syst. Safety*, vol. 91, no. 10-11, pp. 1331-1357, 2006.

# Simulation features

- **Realistic proton beam line**

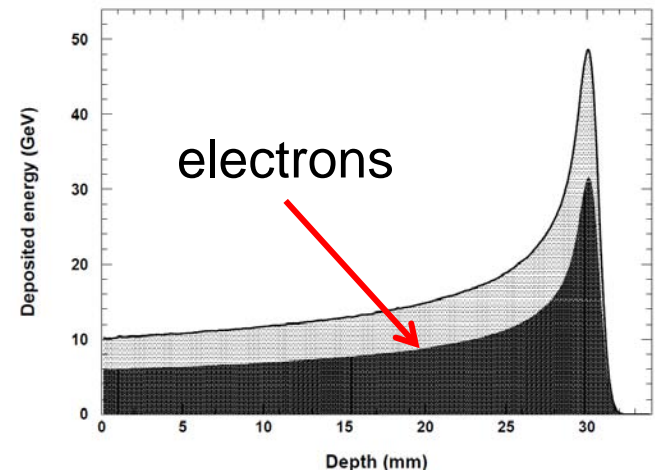
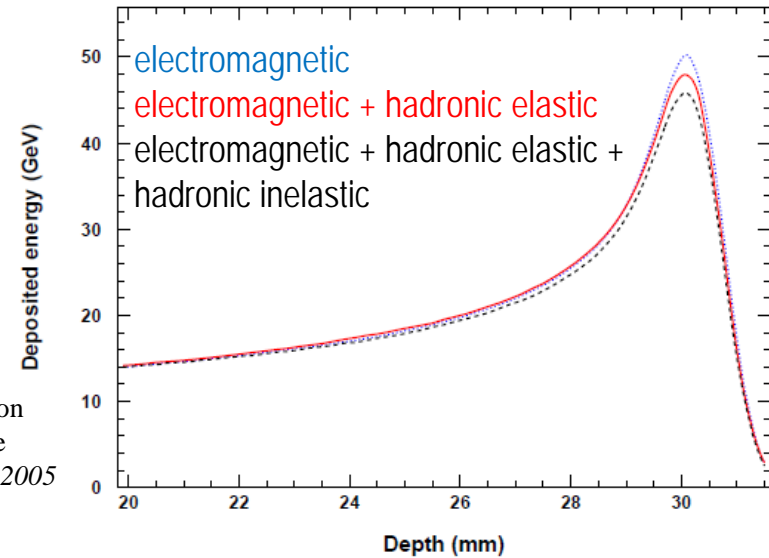
- Geometry from Geant4 *hadrontherapy* advanced example
- G.A.P. Cirrone et al., “Implementation of a New Monte Carlo GEANT4 Simulation Tool for the Development of a Proton Therapy Beam Line and Verification of the Related Dose Distributions”, *IEEE Trans. Nucl. Sci.*, vol. 52, no. 1, pp. 262-265, 2005

- **Water sensitive volume**

- **proton beam:  $E = 63.95$  MeV,  $\sigma_E = 300$  keV**

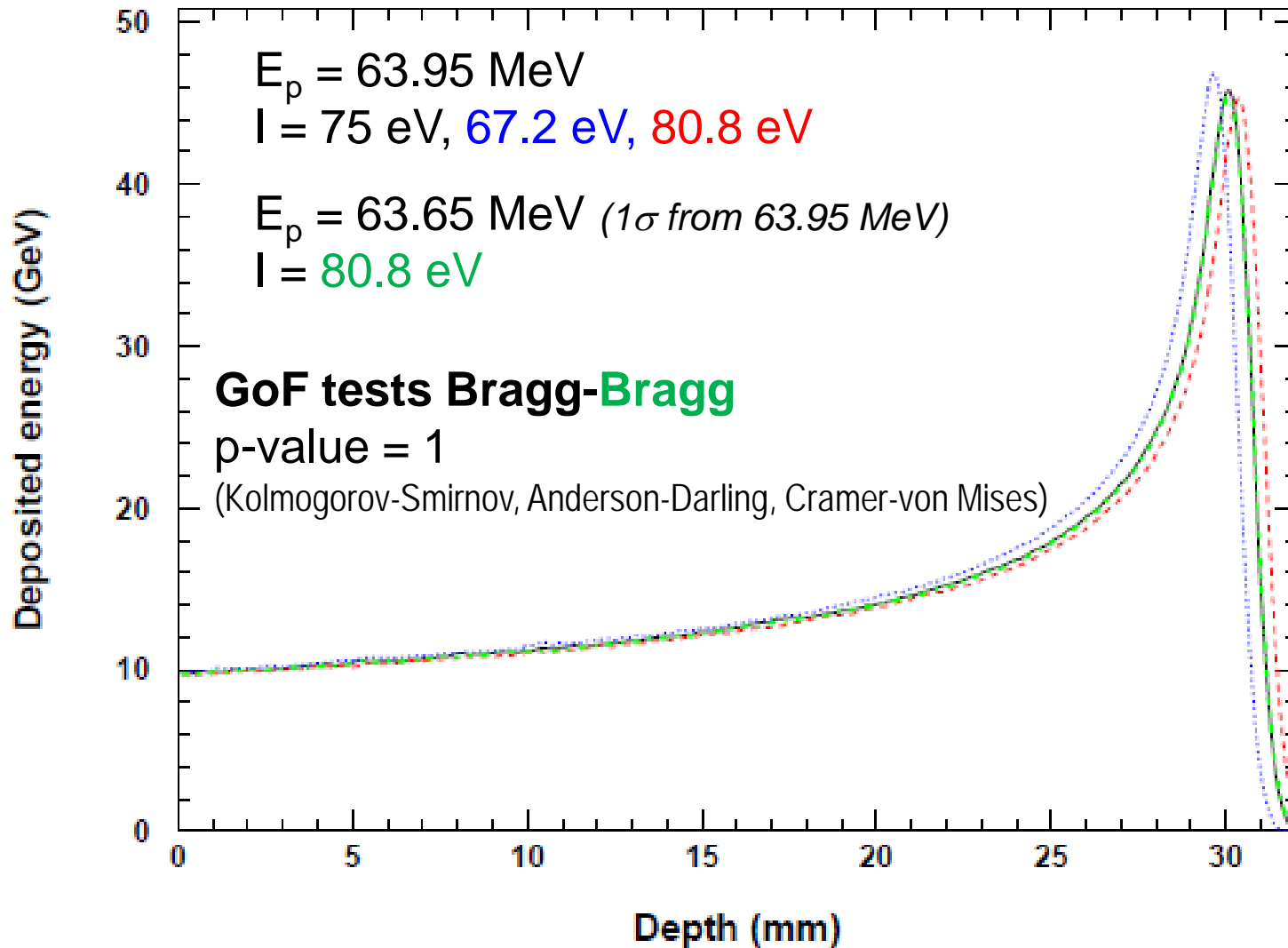
- Physics modeling options in the simulation application configured through a class derived from **G4VModularPhysicsList**

- Geant4 **8.1p02**, **9.1**(ref-04), **9.2p03**, **9.3**, **9.4**



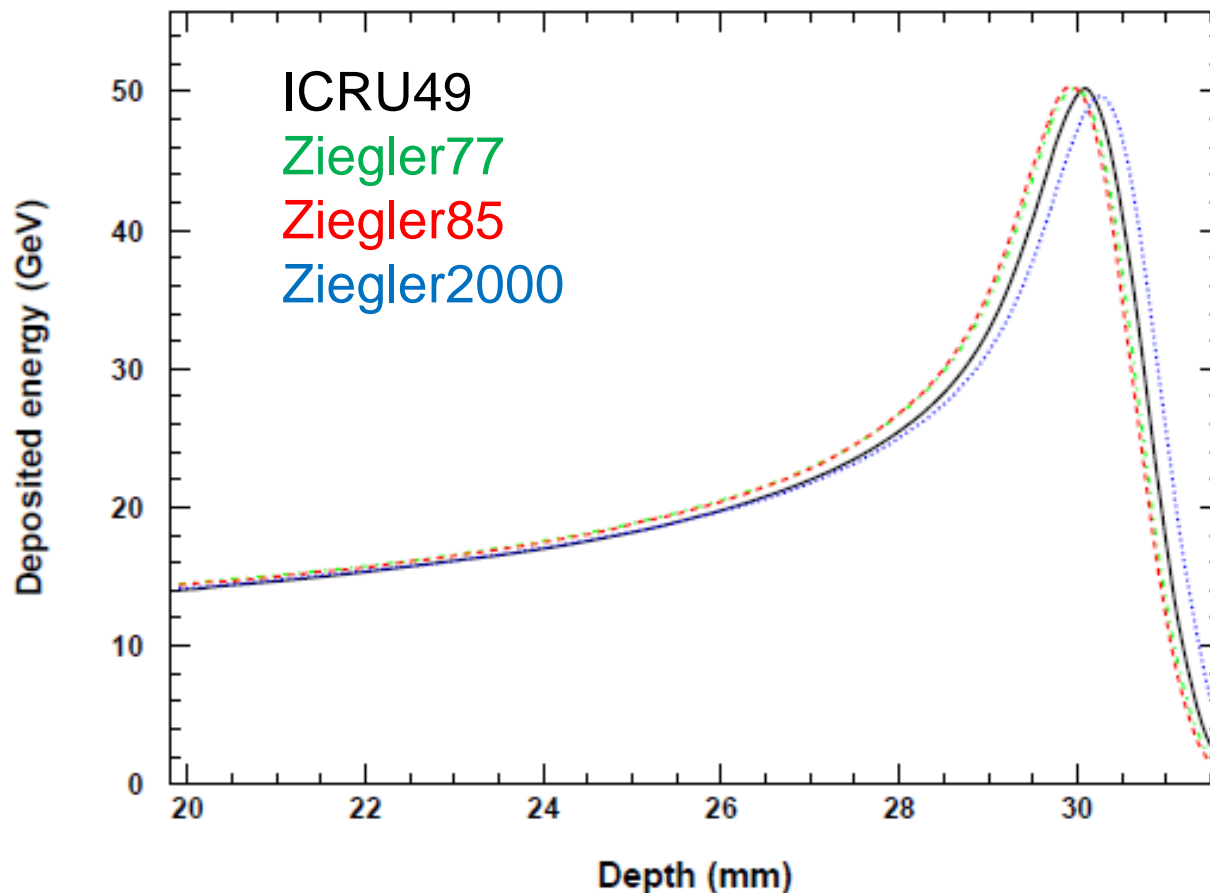
“interval analysis” across different options available in Geant4

# Water mean ionisation potential





# Proton stopping powers

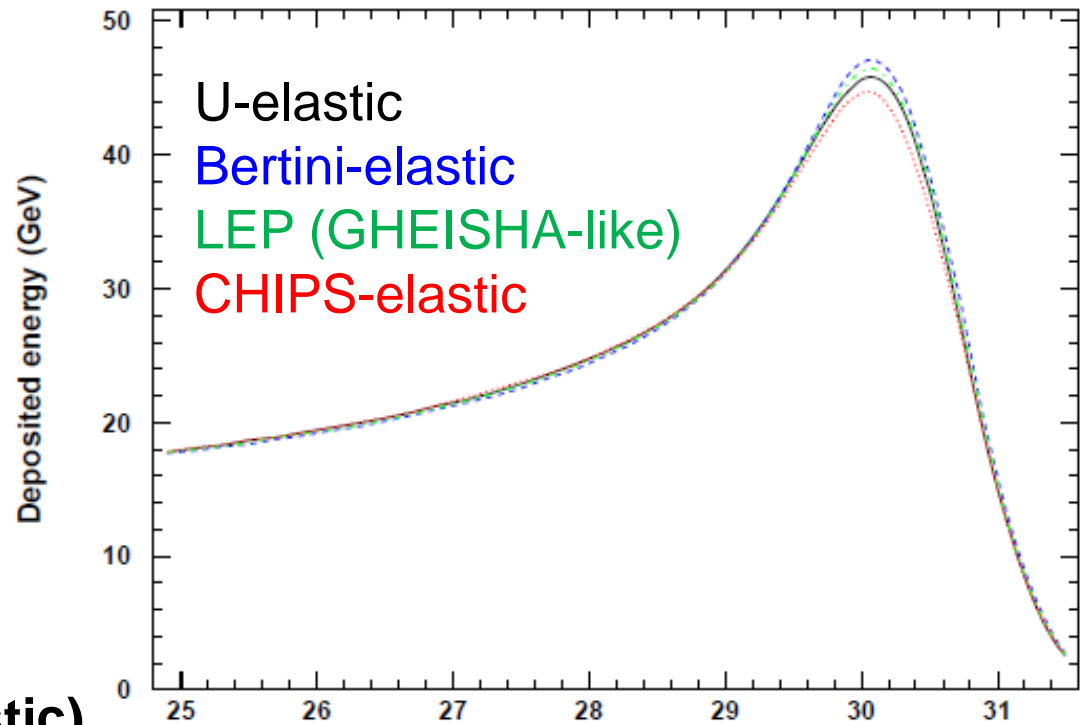


Differences would be masked by typical calibration of simulation input parameters

# Hadronic elastic scattering

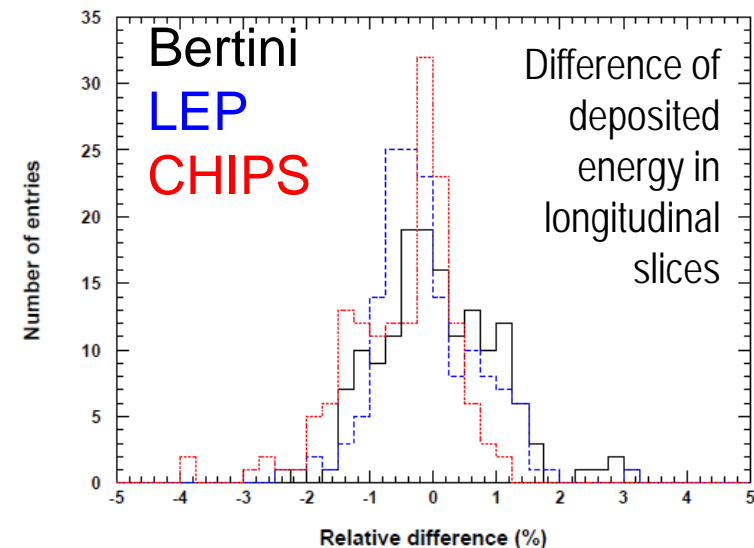
Wald-Wolfowitz test:  
p-value < 0.001

**p-value (reference: U-elastic)**



Depth (mm)

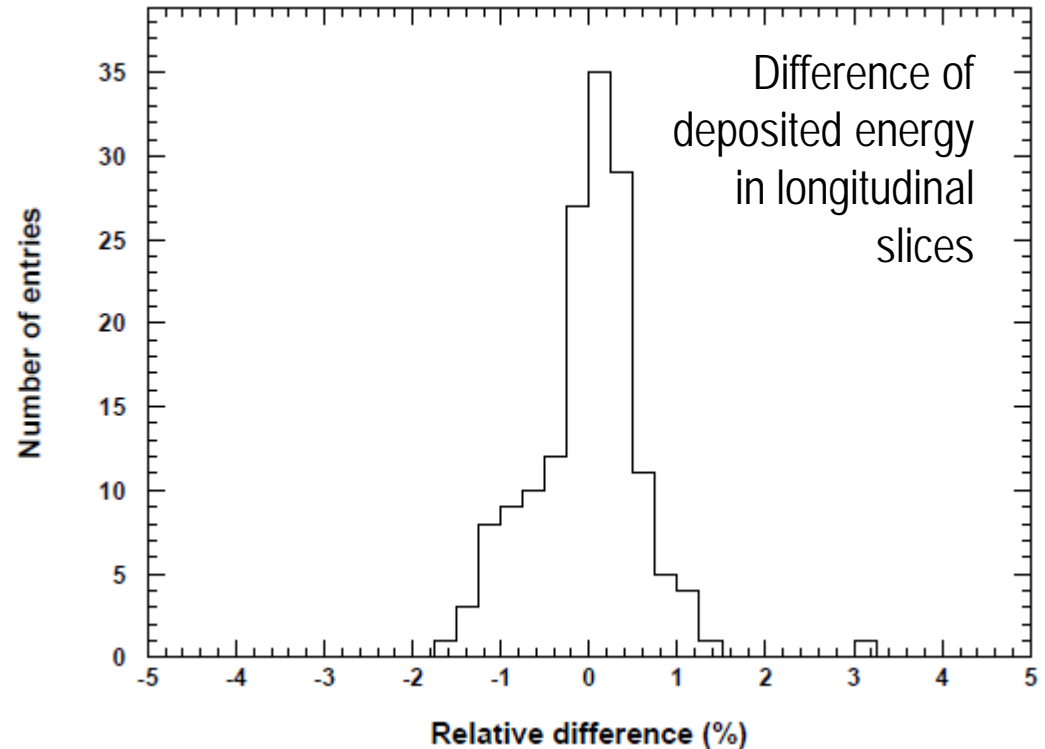
Version	Range	Model	Kolmogorov Smirnov	Anderson Darling	Cramer von Mises
9.3	Whole	Bertini	1	1	1
		LEP	1	1	1
		CHIPS	0.997	0.997	0.999
	Left branch	Bertini	1	1	1
		LEP	1	1	1
		CHIPS	0.996	0.982	0.999
Right branch	Bertini	1	0.986	0.972	
	LEP	1	0.986	0.972	
	CHIPS	1	0.986	0.972	
9.1	Left branch	Bertini	1	1	1
		LEP	0.996	0.989	1
8.1	Left branch	Bertini	1	0.999	0.997
		LEP	1	1	1



# Hadronic inelastic cross sections

- **GHEISHA-like**
- **Wellisch & Axen**

Bragg peak profiles  
**p-value > 0.9**  
(Kolmogorov-Smirnov,  
Anderson-Darling,  
Cramer-von Mises)



**99% confidence interval** for inelastic scattering occurrences in water  
(Wellisch & Axen cross sections): **1688-1849**

Occurrences with **GHEISHA-like** cross sections: **1654**

# Hadronic inelastic scattering models

**p-value (reference: Precompound)**

No visible difference in  
Bragg peak profiles

## Wald-Wolfowitz test

**p-value < 0.001**

for all model options  
except

**p-value = 0.360**

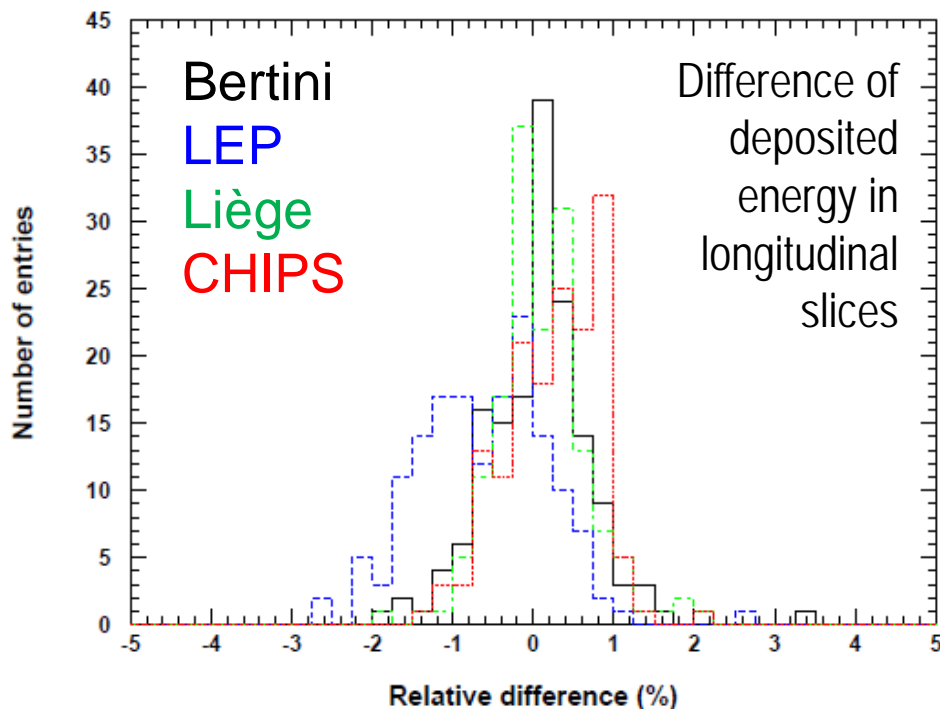
for Liège cascade

Geant4 version	Test range	Hadronic model	Kolmogorov Smirnov	Anderson Darling	Cramer von Mises
9.3	Whole	Bertini	1	1	1
		LEP	0.954	0.988	0.984
		Liège	1	1	1
		CHIPS	1	1	1
		GEM	1	1	1
		Fermi break-up	1	1	1
		Binary	0.954	0.938	0.973
	Left branch	Bertini	1	1	1
		LEP	0.945	0.961	0.979
		Liège	1	1	1
		CHIPS	1	1	1
		GEM	1	1	1
		Fermi break-up	1	1	1
		Binary	0.945	0.858	0.962
	Right branch	Bertini	1	0.986	0.972
		LEP	1	0.986	0.972
		Liège	1	0.986	0.972
		CHIPS	1	0.986	0.972
		GEM	1	0.986	0.972
		Fermi break-up	1	0.986	0.972
		Binary	1	0.986	0.972
9.1	Left branch	Bertini	0.981	0.901	0.980
		LEP	0.945	0.949	0.937
8.1	Left branch	Bertini	1	1	1
		LEP	0.996	0.814	0.847

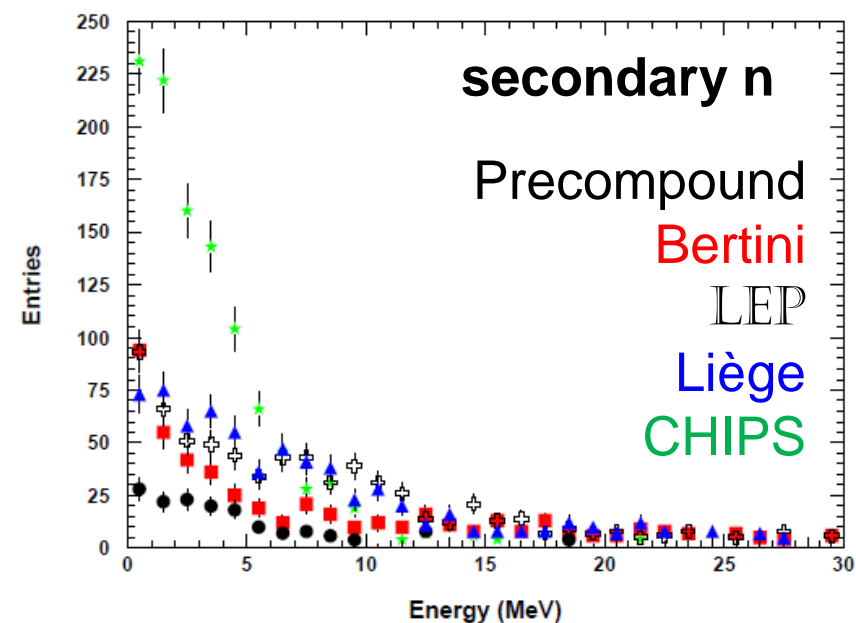
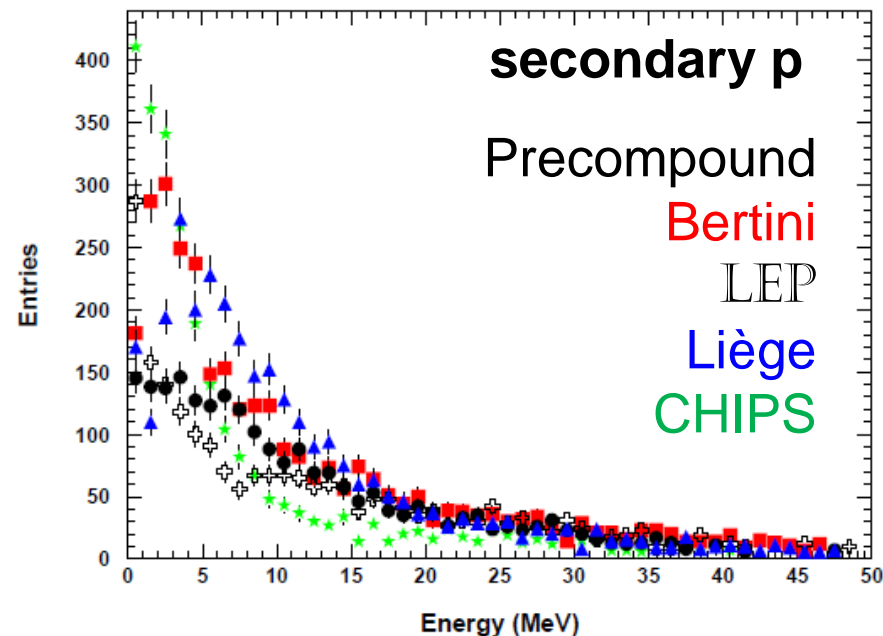
**preequilibrium =  
no preequilibrium**

# Hadronic inelastic differences

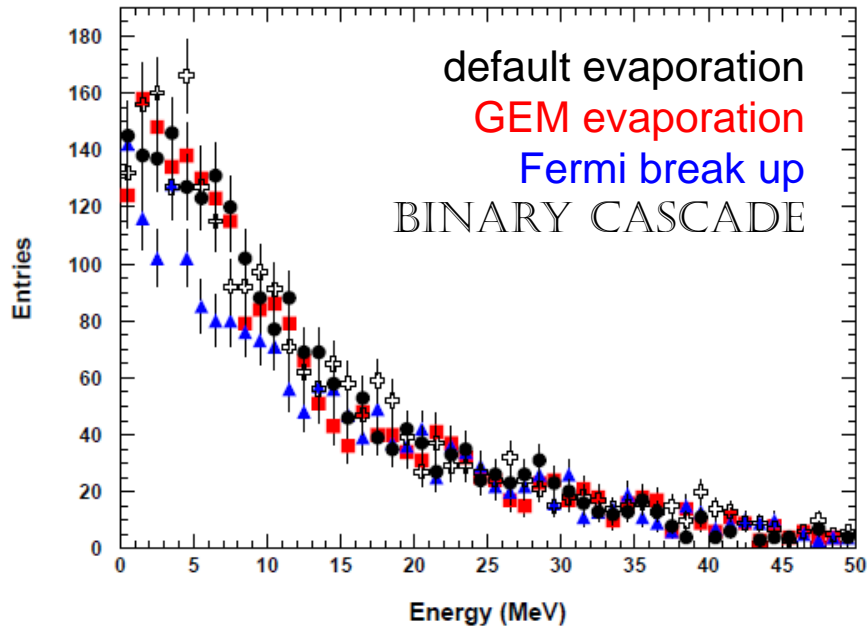
reference: Precompound



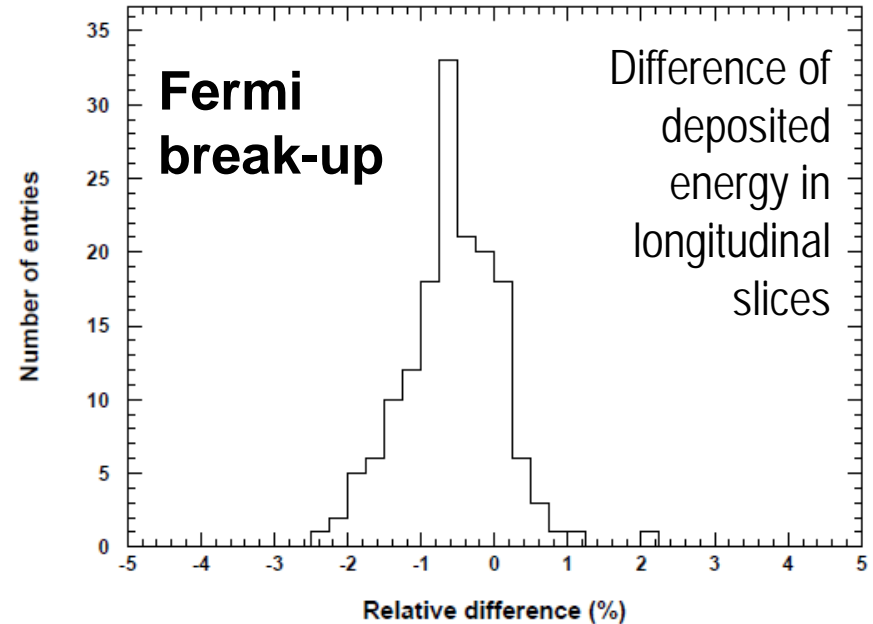
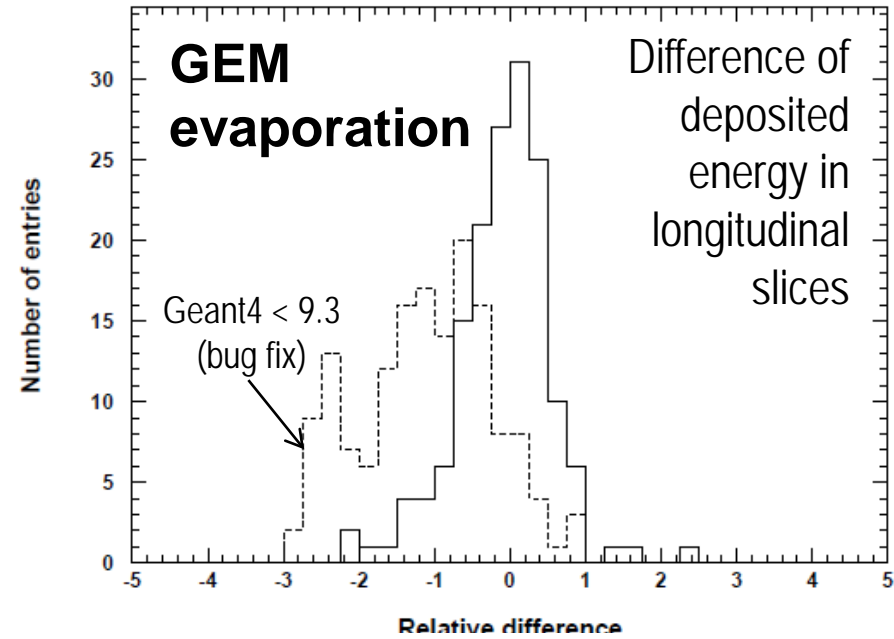
Wald-Wolfowitz test:  $p\text{-value} < 0.001$



# Nuclear deexcitation

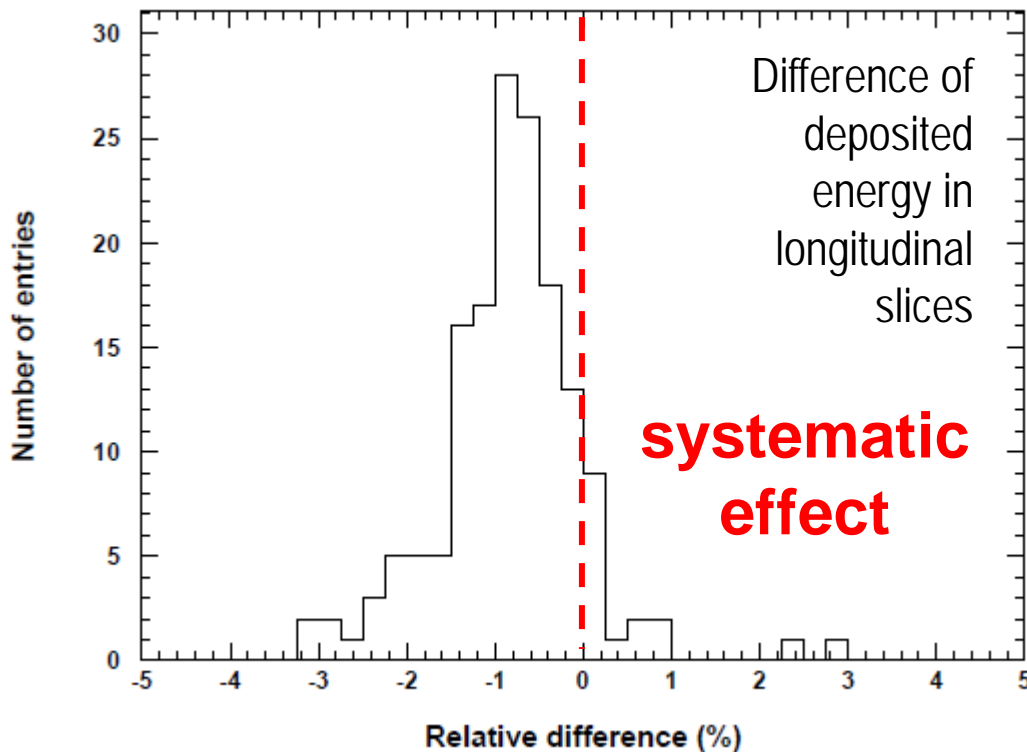


reference: default Evaporation



# Cascade-preequilibrium

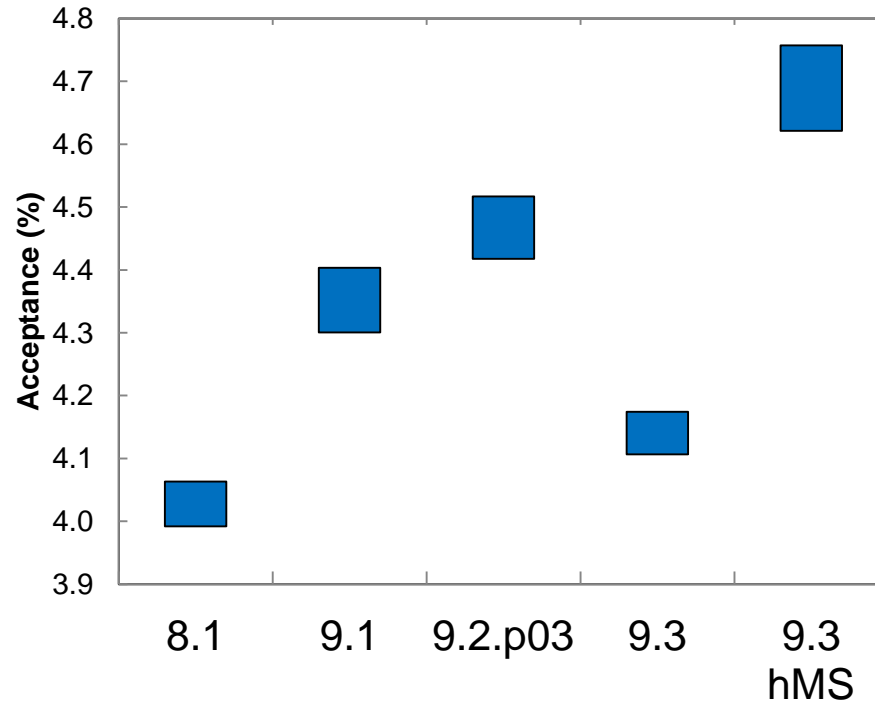
**Precompound** model activated through **Binary Cascade** w.r.t. standalone Precompound model



Transition between intranuclear cascade and preequilibrium determined by **empirical considerations**

In Geant4 Binary Cascade model cascading continues as long as there are particles above a 70 MeV kinetic energy threshold (along with other conditions required by the algorithm)

# Some get lost on the way...



95%  
confidence  
intervals

July  
2006

December  
2009

Calibration: 50 and 200 GeV



# Multiple scattering

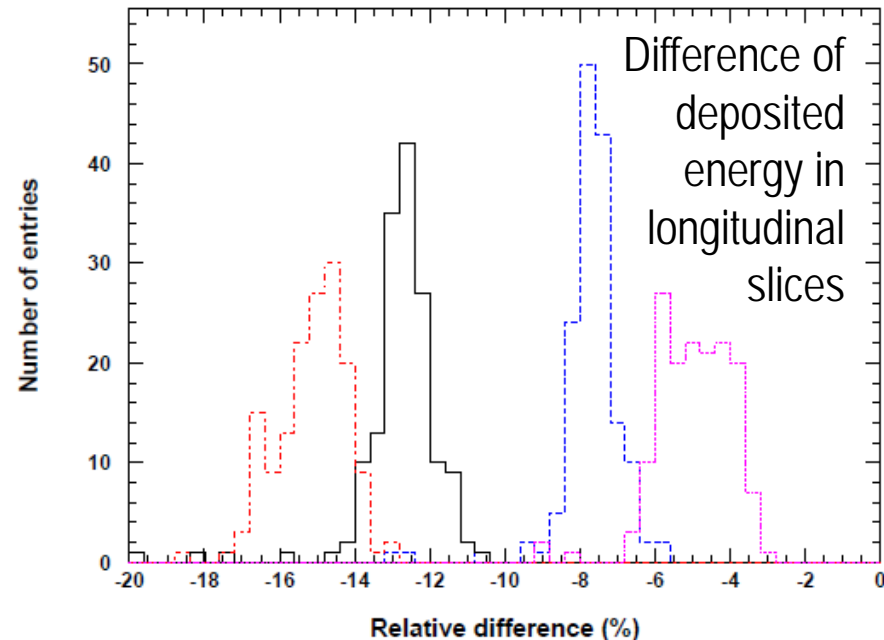
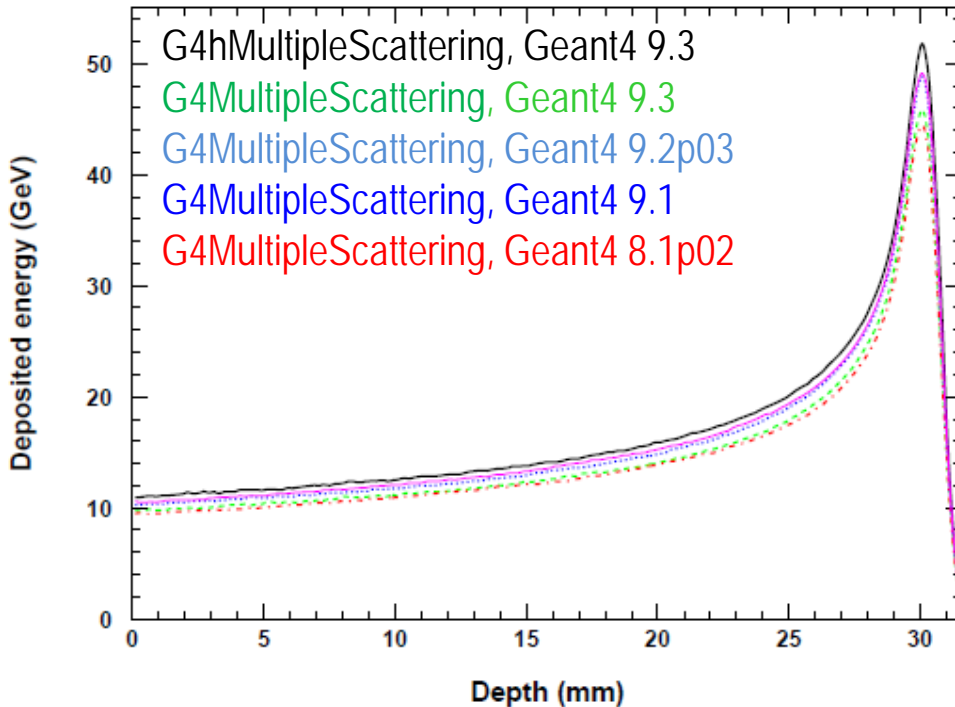
G4MultipleScattering  
G4hMultipleScattering

	RangeFactor	StepLimit	LatDisplacement	skin	geomFactor	Model
<b>8.1</b>	0.02	1				UrbanMSC
<b>9.1</b>	0.02	1	1	0	2.5	UrbanMSC
<b>9.2p0.3</b>	0.02	1	1	3	2.5	UrbanMSC
<b>9.3</b>	0.04	1	1	3	2.5	UrbanMsc92
<b>9.3 hMS</b>	0.2	0	1	3	2.5	UrbanMsc90

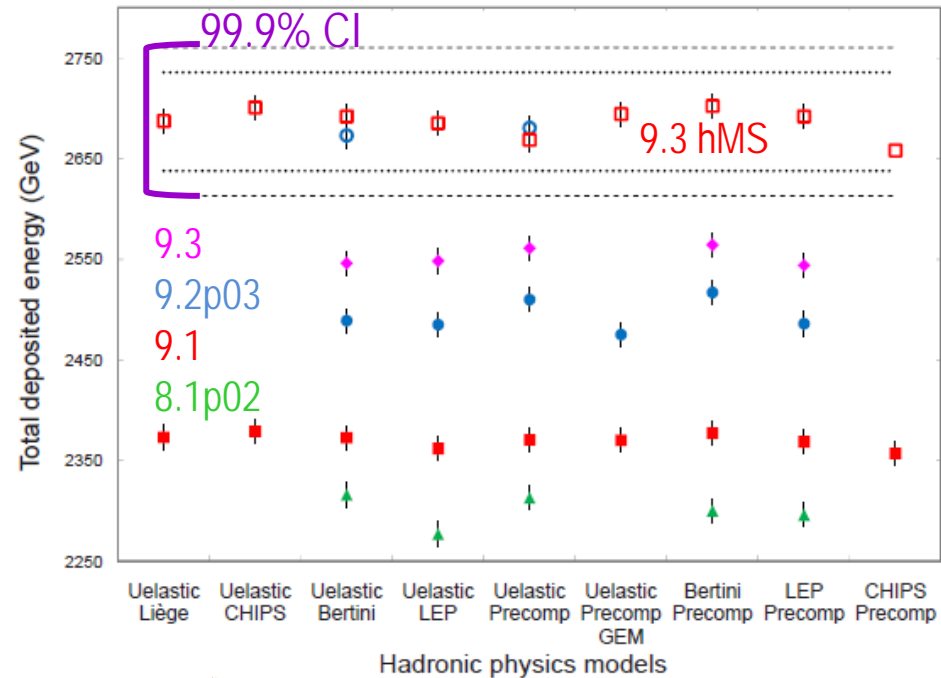
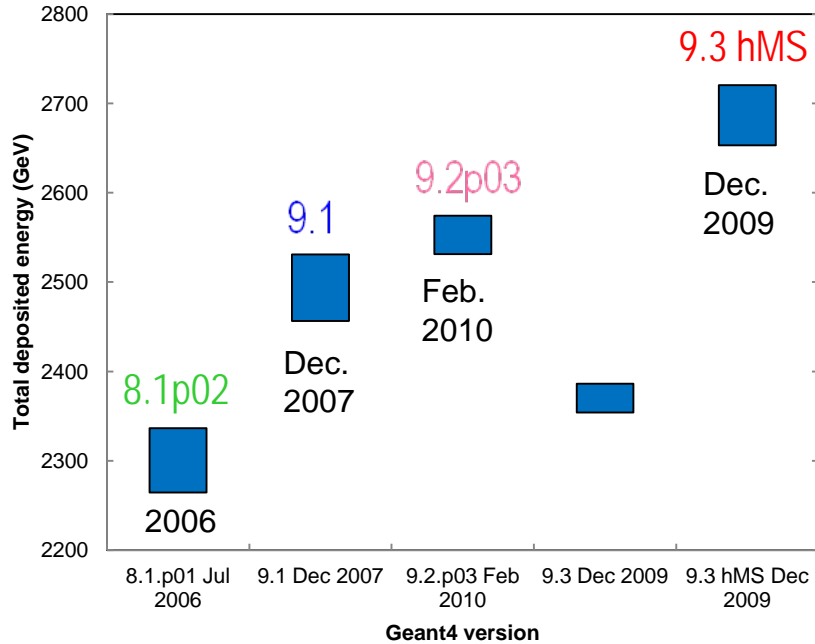
Reference:

Geant4 9.3 G4hMultipleScattering

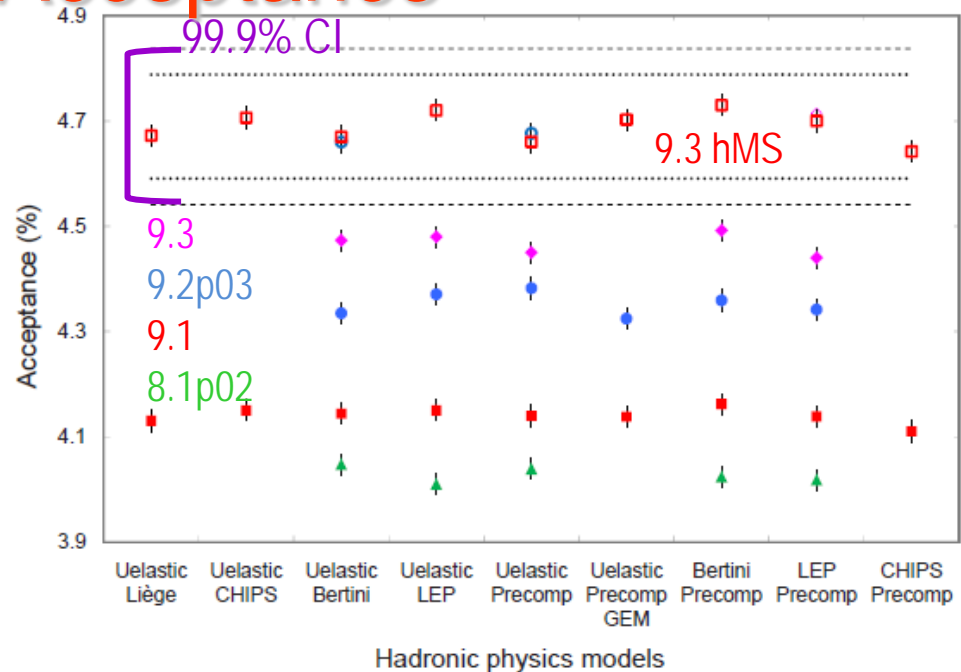
Difference: G4MultipleScattering in Geant4  
9.3 9.1 9.2p03 8.1p02



# Total deposited energy



# Acceptance



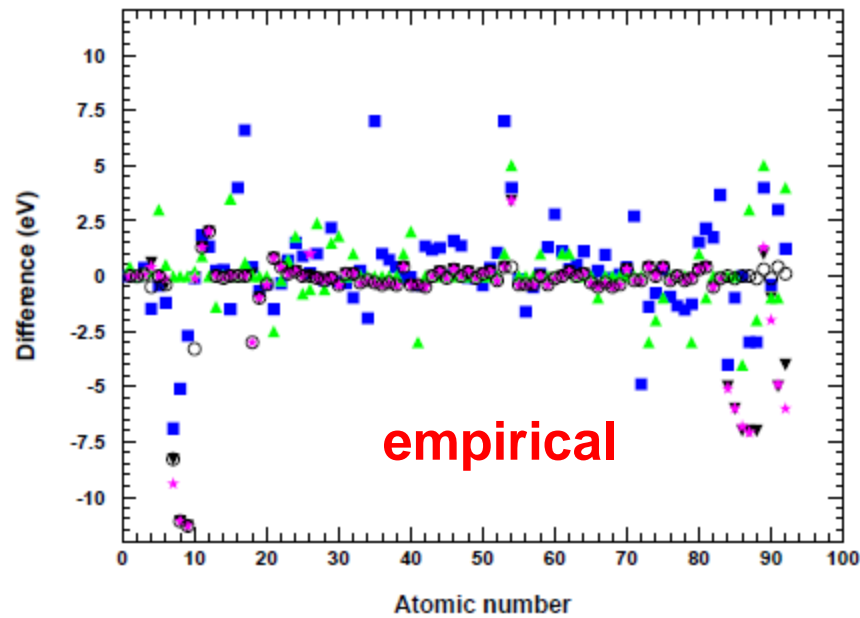
Equivalent across different hadronic models in the same release  
Differences across releases

# How well do we know basic physics parameters in the simulation?

## Atomic electron binding energies

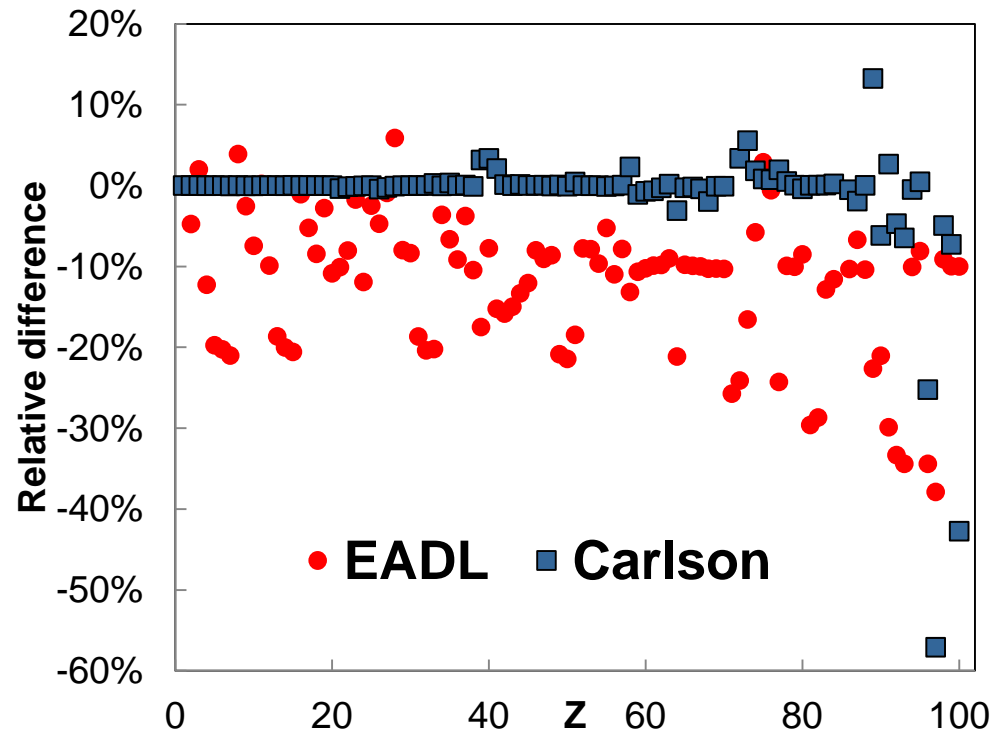
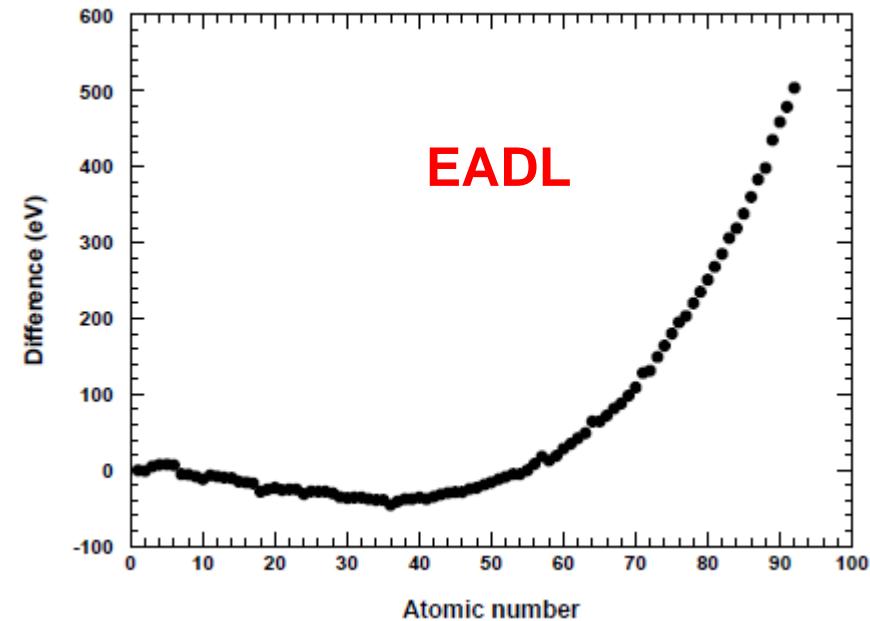
<b>EGS4 EGSnrc</b>	<b>Tol 1978</b> ( <i>Shirley-Uppsala</i> )		
<b>EGS5</b>	<b>Tol 1996</b> ( <i>Larkins(Sevier 1972)</i> )		
<b>FLUKA</b>	<b>Unknown</b>		
<b>GEANT 3</b>	<b>Modified Bearden&amp;Burr</b>		
<b>Geant4</b>	<b>EADL, Carlson+Williams, Tol 1978,</b> (Bearden&Burr)		
<b>MCNP</b>	<b>Carlson</b>		
<b>Penelope</b>	<b>Carlson, Tol 1978</b>		
<b>GUIPIX</b>	<b>Sevier 1979</b>	<b>X-ray Data Booklet</b>	<b>Williams</b>
<b>ISICS</b>	<b>Bearden &amp; Burr, Williams</b>		

# Are they different?



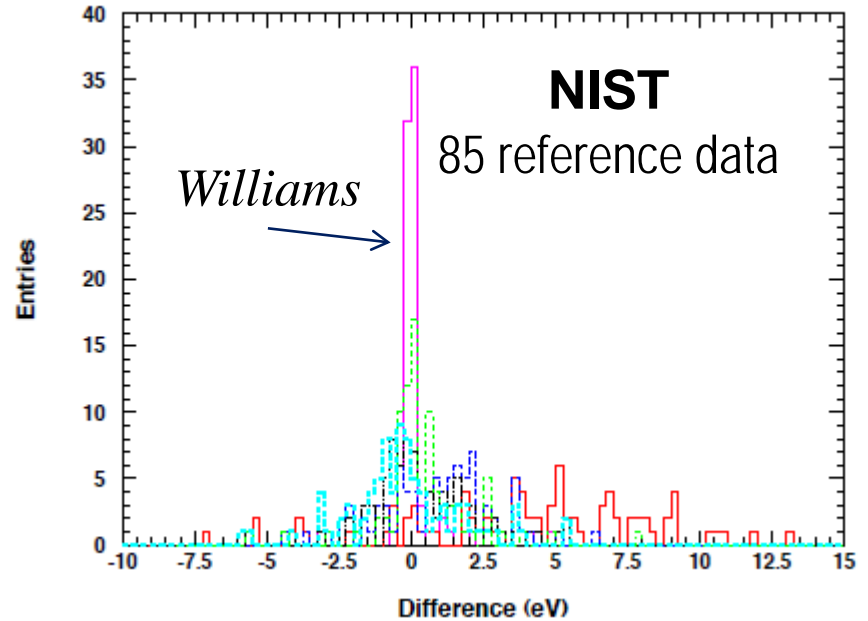
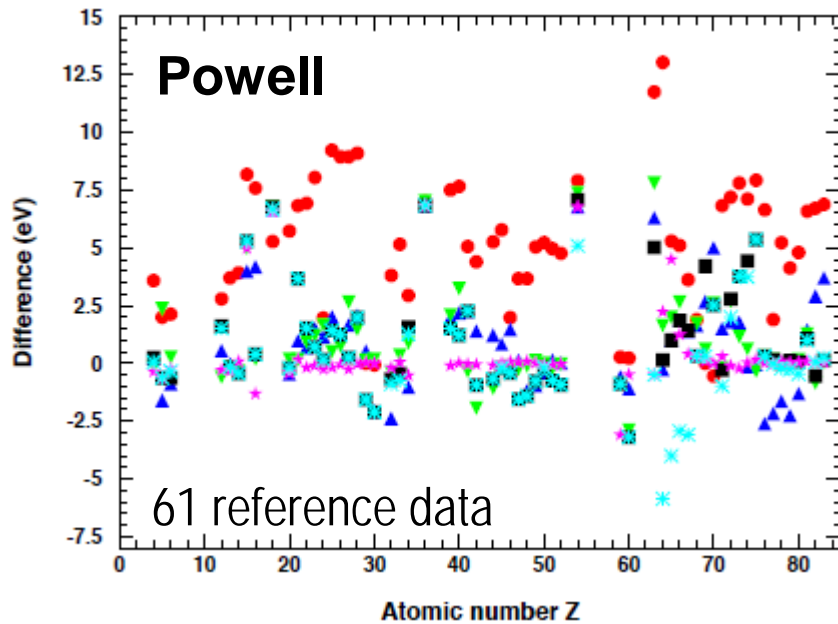
**Ionisation energies**  
(AKA 1<sup>st</sup> ionisation potential)  
w.r.t. NIST experimental

**K-shell (w.r.t. Williams)**



# Direct validation

Whole periodic table: ~1500 shells



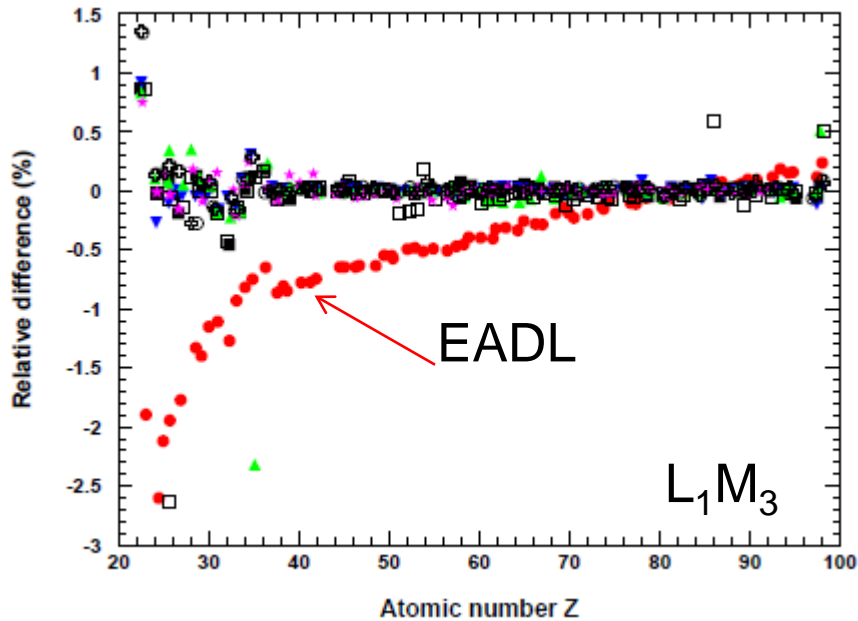
## ● Systematics

- Experimental configuration of measurement
- Chemical and physical effects (*not accounted for in experimental error*)

## ● Usual GoF tests fail

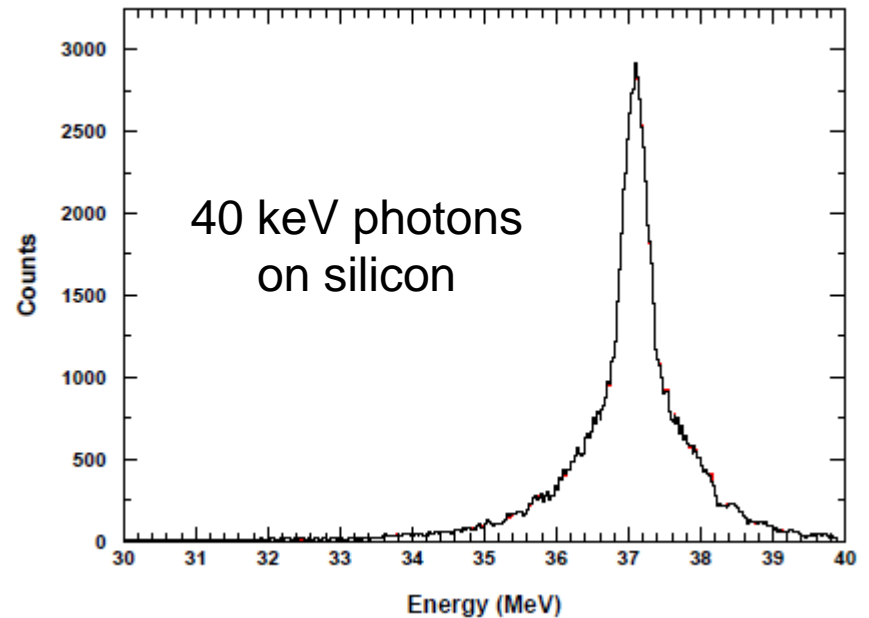
- Underestimated uncertainties, small differences

# Effects on simulation



## X-ray energies

Comparison with  
experimental data  
(review by *Deslattes et al.*)



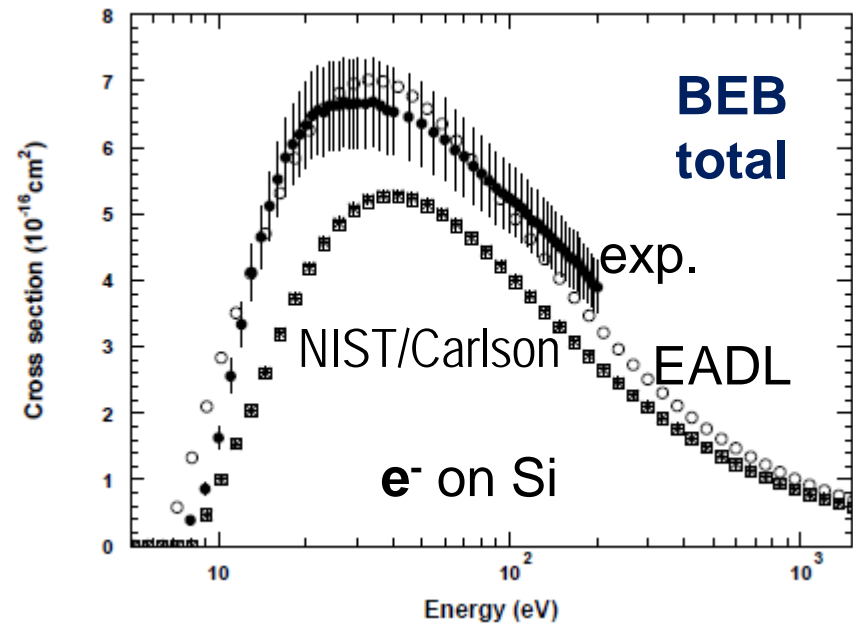
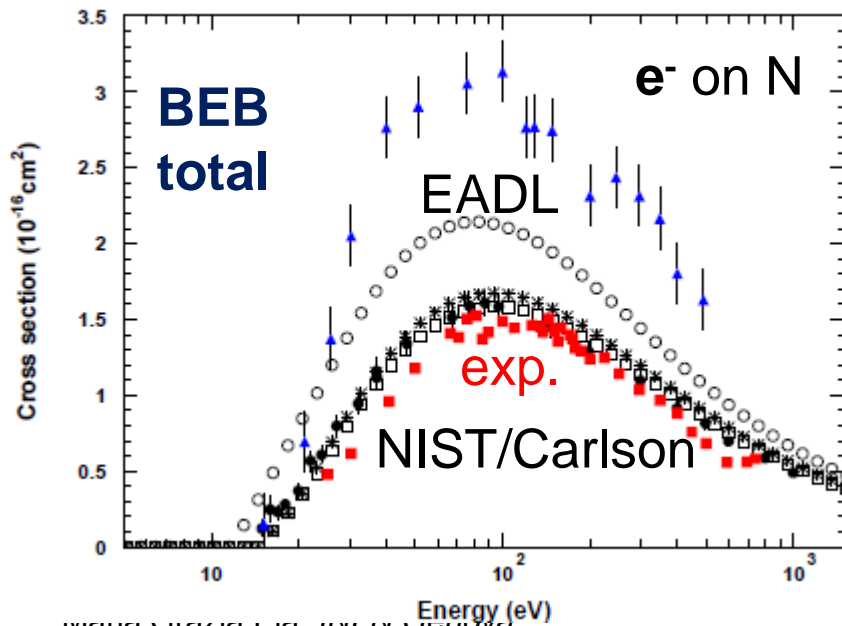
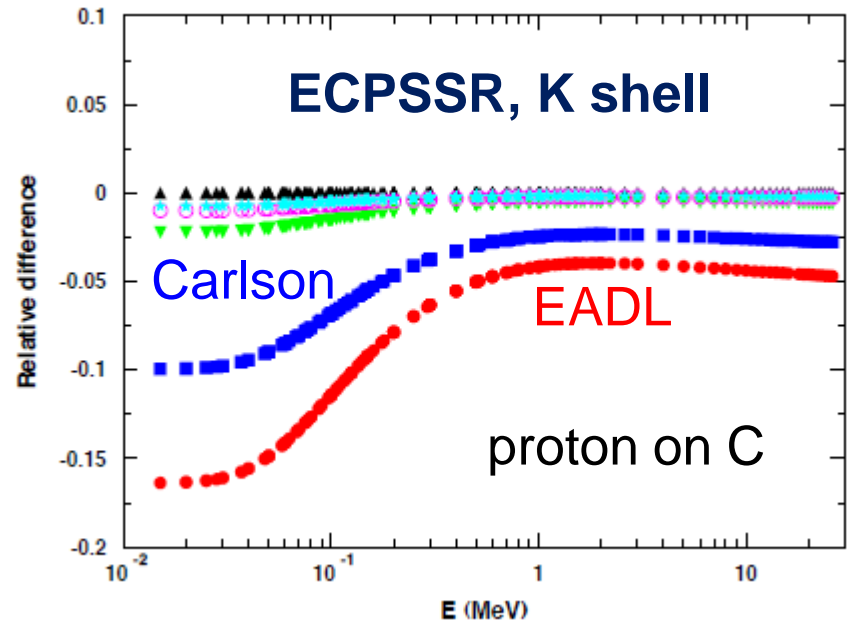
## Compton scattering

Doppler broadened  
photon spectrum

# Ionisation cross sections

Comparison with experimental measurements

*Large scale statistical analysis*



# Benefits of quantifying uncertainties

- Epistemic uncertainties are **reducible**
  - Can be reduced or suppressed by extending knowledge
  - New experimental measurements
- Uncertainty quantification gives us guidance about
  - **What to measure**
  - What experimental **precision** is needed/adequate
  - **Priorities**: which uncertainties generate the worst systematic effects
- Measurements are not always practically possible
  - Uncertainty quantification to control systematics

Especially important in critical applications



# Conclusions

- Evaluation of **systematic effects** associated with **epistemic uncertainties**
  - Sensitivity analysis (~interval analysis)
  - More refined methods: Dempster-Shafer
  - Methods specific to Monte Carlo simulation?
- **Complementary statistical methods** contribute to identify and quantify effects
  - Qualitative appraisal is not adequate
- Epistemic uncertainties are **reducible**
  - Can be reduced or suppressed by extending knowledge
  - New experimental measurements
- Uncertainty quantification gives us **guidance** about
  - What to measure
  - What experimental precision is needed/adequate
  - Priorities: which uncertainties generate the worst systematic effects

The impact of epistemic uncertainties depends on the **experimental** application environment

# INFN UQ

- Project on **Geant4 uncertainty quantification**
- Discussed this week in INFN Computing Committee
  
- Complementary expertise
  - Geant4 developers
  - Experimental groups
  - Theory
  - Mathematics/statistics
- Collaboration is welcome!