IUPS Virtual Physiological Human Satellite

Discussion Counter-Point:
Public Access to Data and Models is Impractical and Counterproductive

James Bassingthwaighte / Peter Kohl
www.physiome.org / ww.vph-noe.eu

Ritsumeikan University, 25 July 2009

Reproducibility is overrated

- *Better ideas* will emerge when other investigators attack a model (hypothesis) and attempt to improve on it.
- *Redevelopment* is the best way to penetrate a model or hypothesis – i.e. no waste of effort!

  This

  - is the real test of reproducibility of the original authors' views

  - really validates it by doing it totally independently

  - makes it less likely to incorporate the original's errors

Copy-cat errors are like mutations in a phylogenetic sequence, so it's better to start at the beginning each time.

Prevent bad science: misuse & hype

- *Secondary users* often don't take the time to understand the basics of a model developed without their input.
- *Redevelopment* supports removal of inconsistencies with data and sparks alternative model development.
- *Opportunistic followers* will generate **hype** that is detrimental to the field as they

  - won't understand implicit or explicit assumptions

  - won't pay attention to limitations of conditions under which the model has reasonably correct behavior.

All models are "wrong" and casual users won't realise problems and restrictions.

Practicalities: extra work

- *Academic realities:* grants favour novel ideas, not tidying up of established findings. Instead of tidying up, I should be creative.

- *Journal review*: restricted resources (time, expertise, hardware, funding) prevent peers from reproducing methods, whether experimental or computational.

- *Repositories:* existing experience (Physiome, CellML and Biomodels) shows that published models, even provided by willing authors, need extensive checking and support.

Cook-book writing requires a specific and dedicated effort, outside of and in addition to the basics of academic research.

Practicalities: responsibility

- *Sharing = raising expectations: by sharing you commit to* supporting efforts of others.

- *Poor standards*: most models are developed for a specific purpose, and not by professional developers.

Hobbyist's work should not be distributed – out of professional courtesy!

Practicalities: career suicide

- *Academic evaluation:* is based almost solely on peer-reviewed publication, scored by impact factor of journals. There are no points for model sharing.

- *On the contrary*: making the fruits of year-long model development efforts available will relieve authors of control over and involvement in model application, with the result that they have one paper (in a low-impact maths or computation) journal yet no benefit from subsequent application.

Model sharing = cutting the branch on which developers sit.

Practicalities: lost opportunities

- *Commercial realities:* protection of intellectual property is key to commercial exploitation.

  One way of preventing implementation of new ideas is to make them publicly available: in the absence of protectable IP, no company will take the lead in developing applications that can be copied.

If you want to do god, you have to be evil at times...

And finally – why me?

- *Ethos versus reality:* sharing & caring is fine, but why would I
  - invest time, funds & effort (all of which are in short supply)
  - defer other original work that I could do instead
  - risk career progression
  - surrender intellectual property and
  - endanger commercial opportunity

  when there is no benefit to me?

I need to pay my mortgage.

Challenges

- *Privacy vs. Sharing and (Re-)Using Data*

  Personalised medicine? Collaboration?
- *Application of Population Data to Single Individual*

  Benefit vs cost?
- *IP vs. Sharing Models & Data*

  Patents key for industrial application
- *Standards vs. Freedom and Development*

  Are standards restrictive?
- *Data & Model Sharing: Yes, No, or Maybe?*

  At time of publication everything should be reproducible

  Lobby editors & reviewers to ensure access to models / data / metadata / range of applicability (validated use).

  Validation: what does it mean / how to do it?

  Encourage funding agencies to fund open access?

  Change reward system to acknowledge model sharing?

  Licence / reference to models / public key encryption?

  Indicators of interdisciplinarity → explore their use.