# Where is Neptune?

## Evolution on graphs and vestibular schwannoma

Chay Paterson

University of Manchester

September 2022

# The big picture

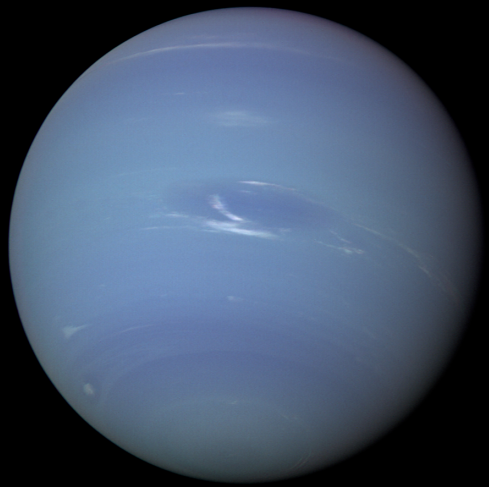
The value of mechanistic models

- Model with specific mechanism
- Exhausted known factors
- Residuals = new factor
- Model predicted **size and location** of unknown planet

[1] J.P. Nichol, "The planet Neptune: an exposition and history" 1849

# The big picture

## The value of mechanistic models

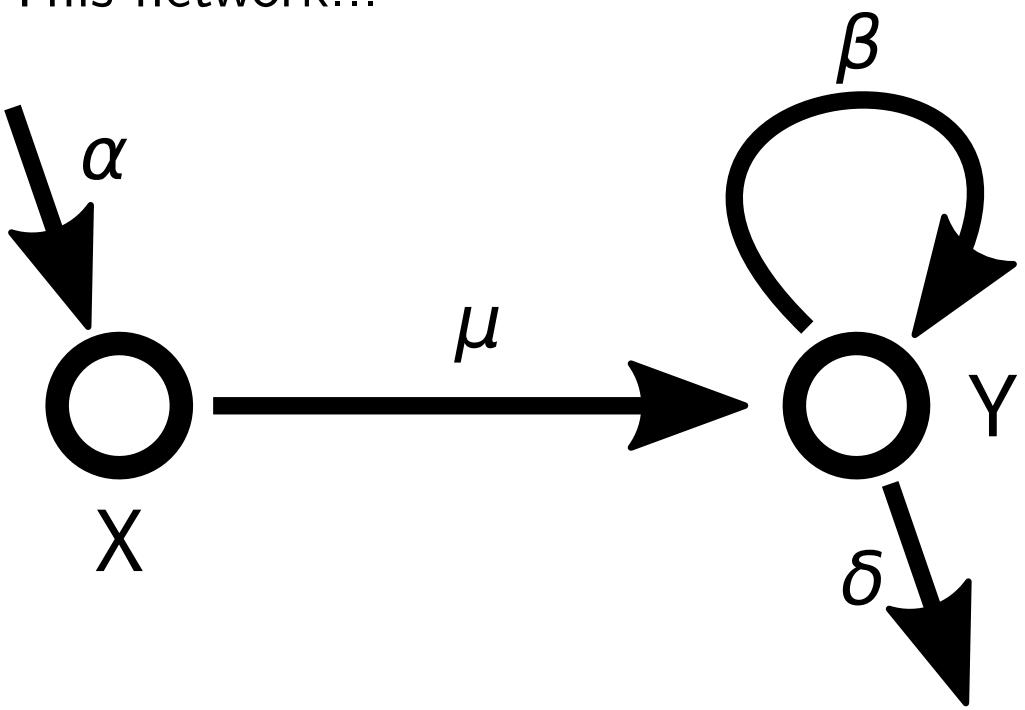


↑ within 1° of predicted location

[1] J.P. Nichol, "The planet Neptune: an exposition and history" 1849

# Some notation

This network...




corresponds to this stochastic process:...

$$\begin{aligned} \emptyset &\to X, && \text{rate } \alpha \\ X &\to X + Y, && \text{rate } \mu X \\ Y &\to Y + Y, && \text{rate } \beta Y \\ Y &\to \emptyset, && \text{rate } \delta Y \end{aligned}$$

and approximately this linear system...

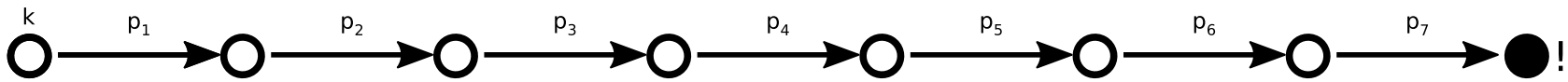
$$\frac{d}{dt}\begin{pmatrix} E[X] \\ E[Y] \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ \mu & \beta - \delta \end{bmatrix} \cdot \begin{pmatrix} E[X] \\ E[Y] \end{pmatrix} + \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$$

Most of our models are linear, high-dimensional and sparse[1]

[1] C. Paterson, I. Bozic, H. Clevers, PNAS 2020; 117(34): 20681-20688

# Multi-stage models

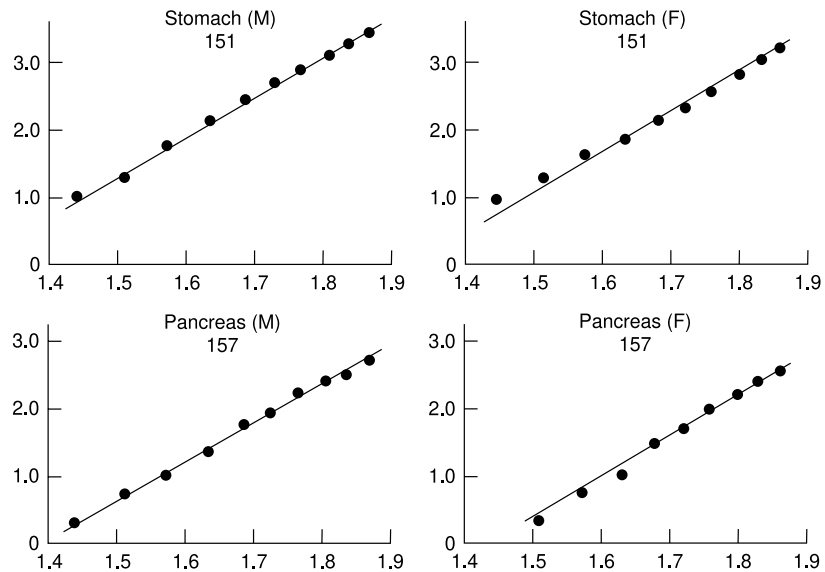
## P. Armitage and R. Doll[2]




$$\downarrow$$

$$P(\text{cancer}) \approx k p_1 p_2 p_3 p_4 p_5 p_6 p_7 \frac{t^7}{7!}$$

$$\Longrightarrow \text{ incidence} \propto \frac{t^6}{6!}$$




[1] P. Armitage and R. Doll, British Journal of Cancer 1954; 8: 1–12

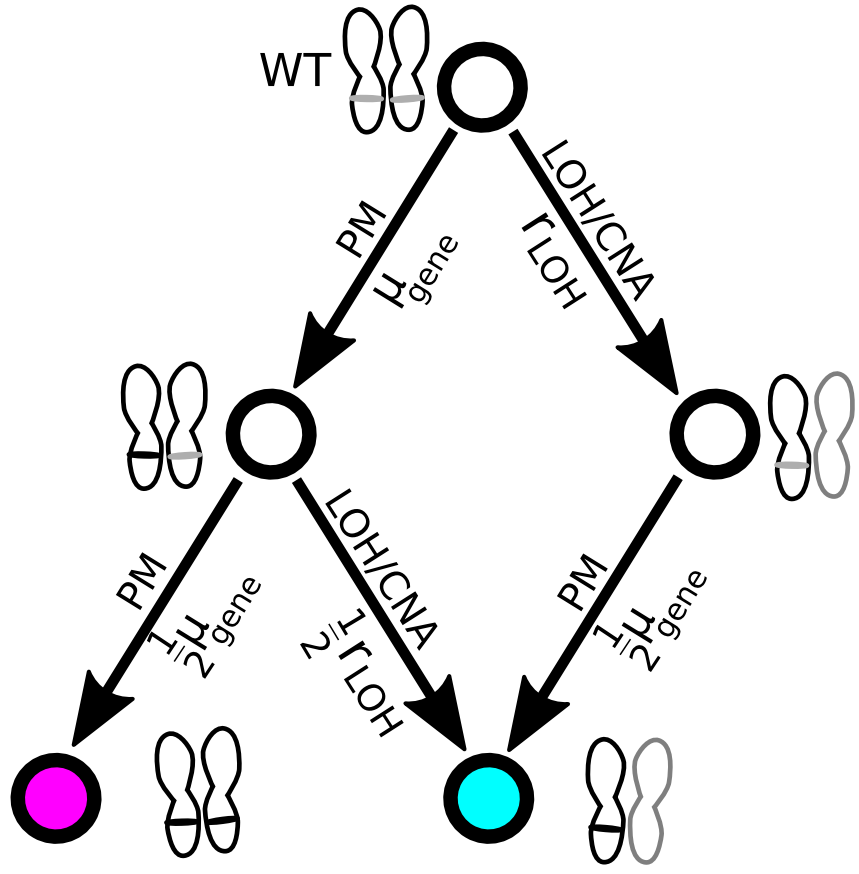
[2] note that $P(t) = 1 - S(t)$, other authors (e.g. Knudson)

# Mechanistic network models

1. Study **specific genes** and mechanisms of interest
2. Fix parameters from sequences and experiments
3. Distinguish different orders of events

**Predict copy number alterations (etc.)**

example model




This gets us the incidence of *specific karyotypes*

# Vestibular schwannoma incidence

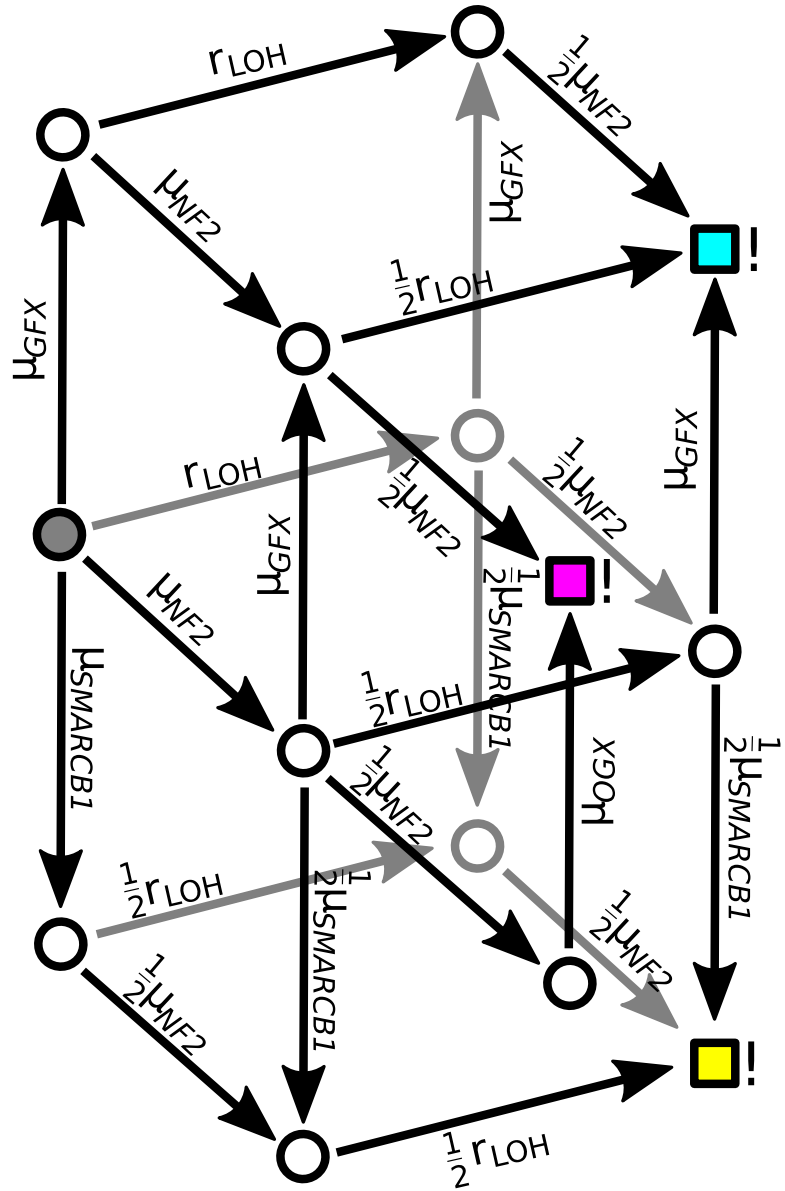
## Our model for sporadic VS

- 3 events[2]
- Include *NF2*, *SMARCB1* and (simplified) linkage
- Add hypothetical oncogene *GFX*

Risk of each subtype looks like

$$P(\blacksquare) \propto \frac{t^3}{3!}$$

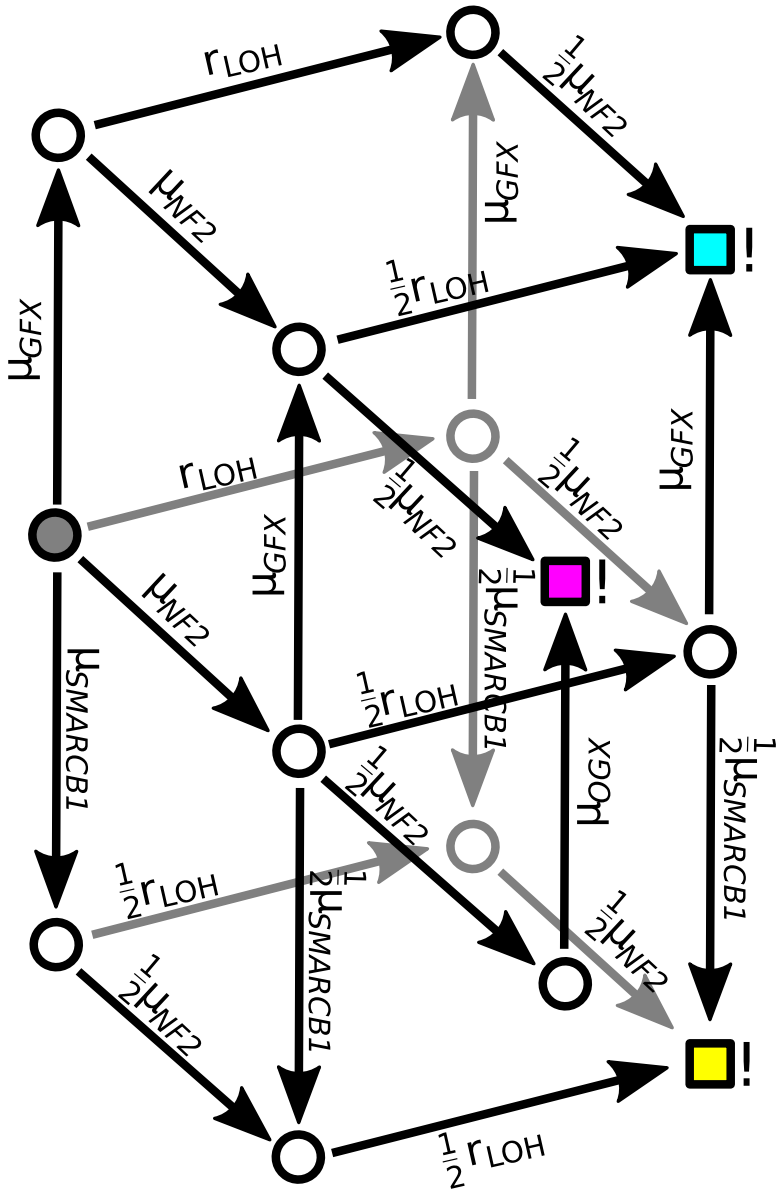
$$P(\blacksquare) \approx N_{WT}\mu_{GFX}r_{LOH}\frac{1}{2}\mu_{NF2}\frac{t^3}{3!}\times 6$$




[1] C Paterson, I Bozic, MJ Smith, X Hoad, DGR Evans, https://doi.org/10.1101/2021.10.03.457528

# Vestibular schwannoma incidence

## Our model for sporadic VS



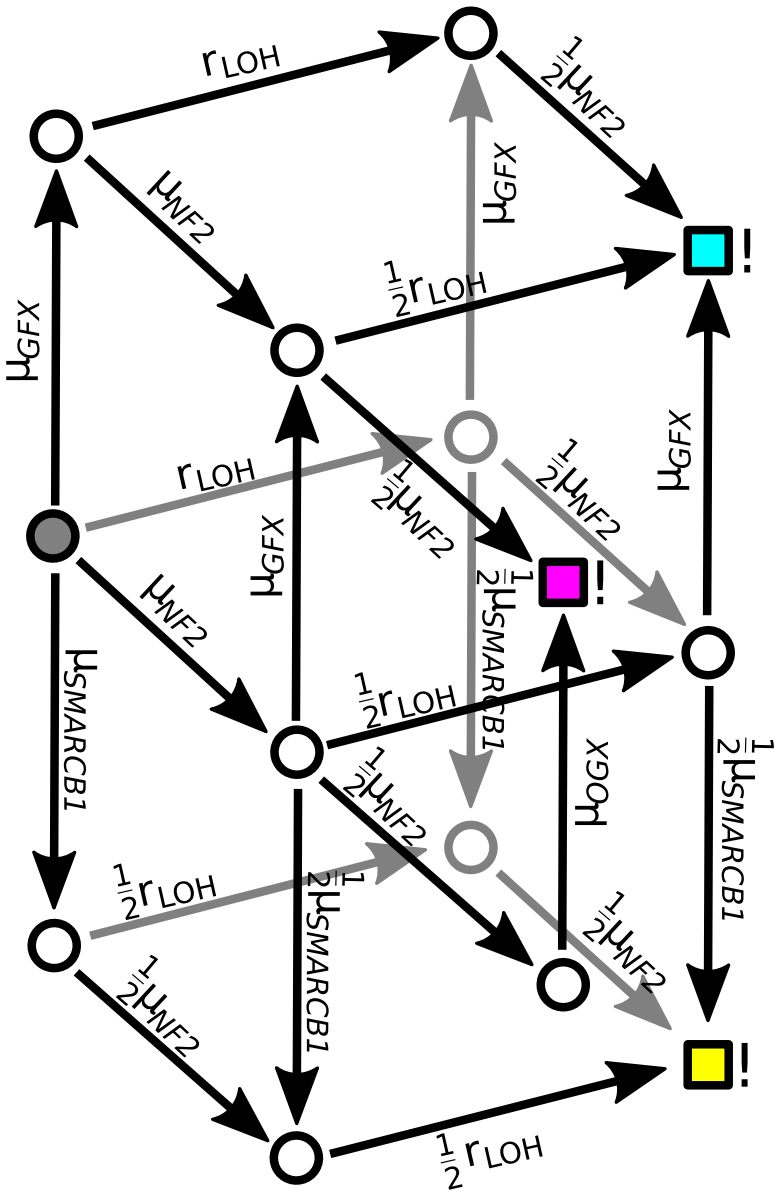

- ■ + ■ = have LOH on 22q
- frequency of LOH = $f_{LOH} = (■+■)/(■+■+■)$
- ■ = *SMARCB1*$^{-/-}$
- frequency of *SMARCB1*$^{-/-}$ = $f_{SMARCB1}$

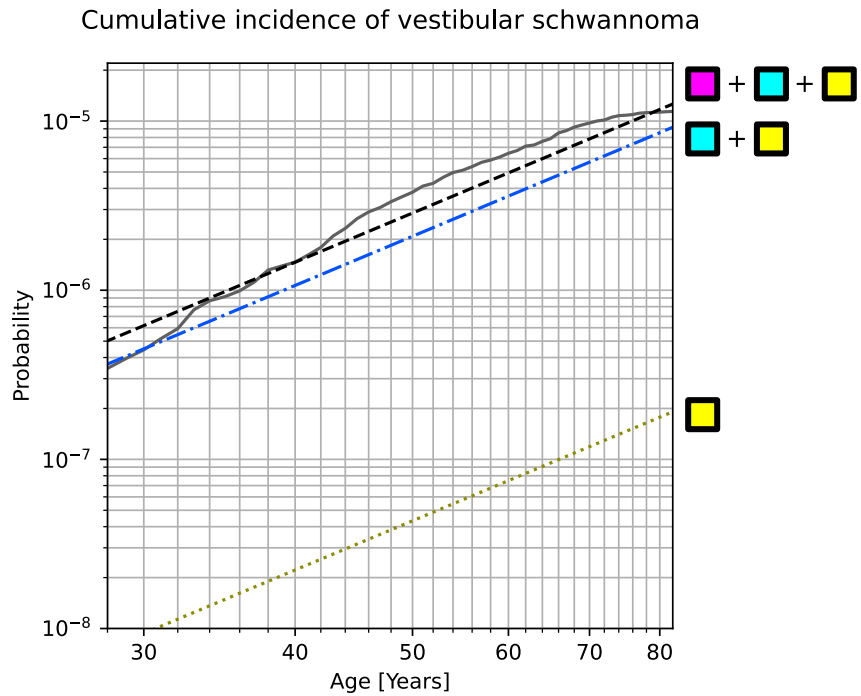
  = ■/(■ + ■ + ■)

Can use these to fix parameters!

# Vestibular schwannoma incidence

## Our model for sporadic VS

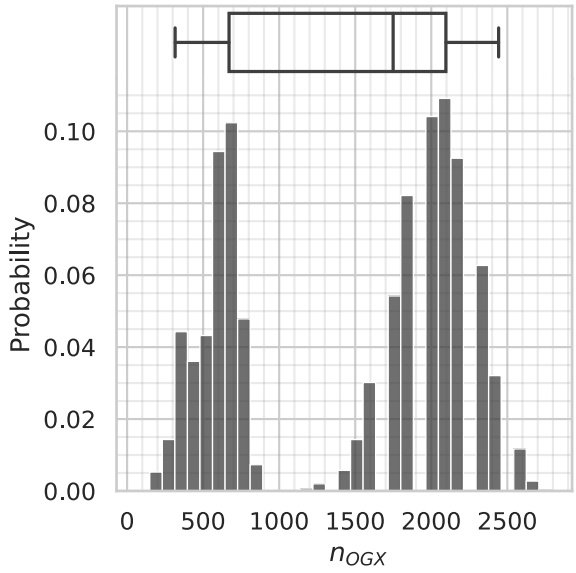


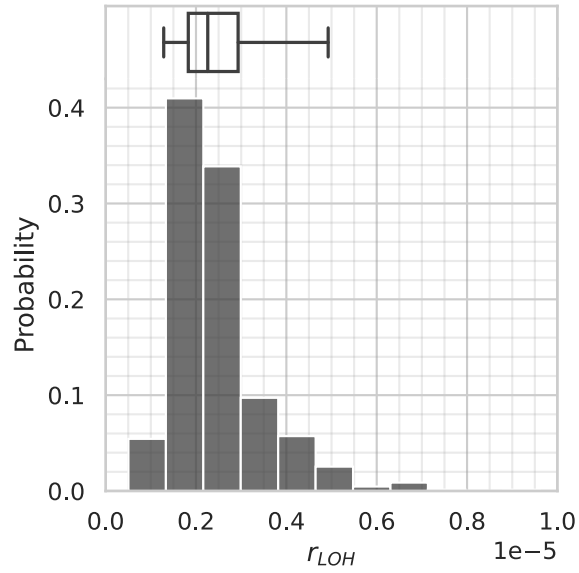



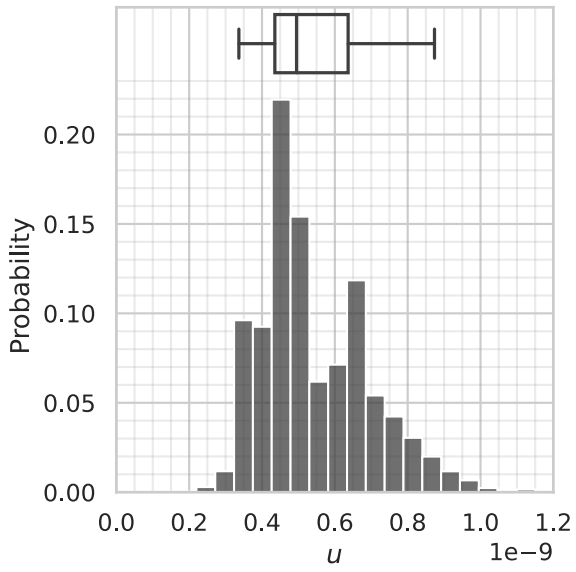

[1] C Paterson, I Bozic, MJ Smith, X Hoad, DGR Evans, https://doi.org/10.1101/2021.10.03.457528

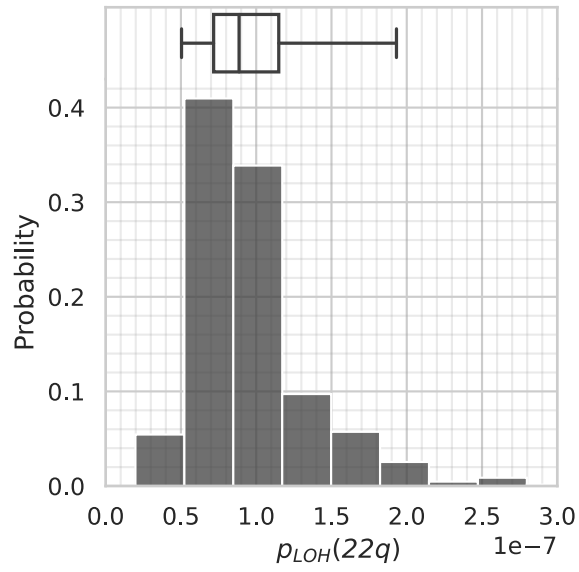
# Vestibular schwannoma incidence

## New parameter estimates



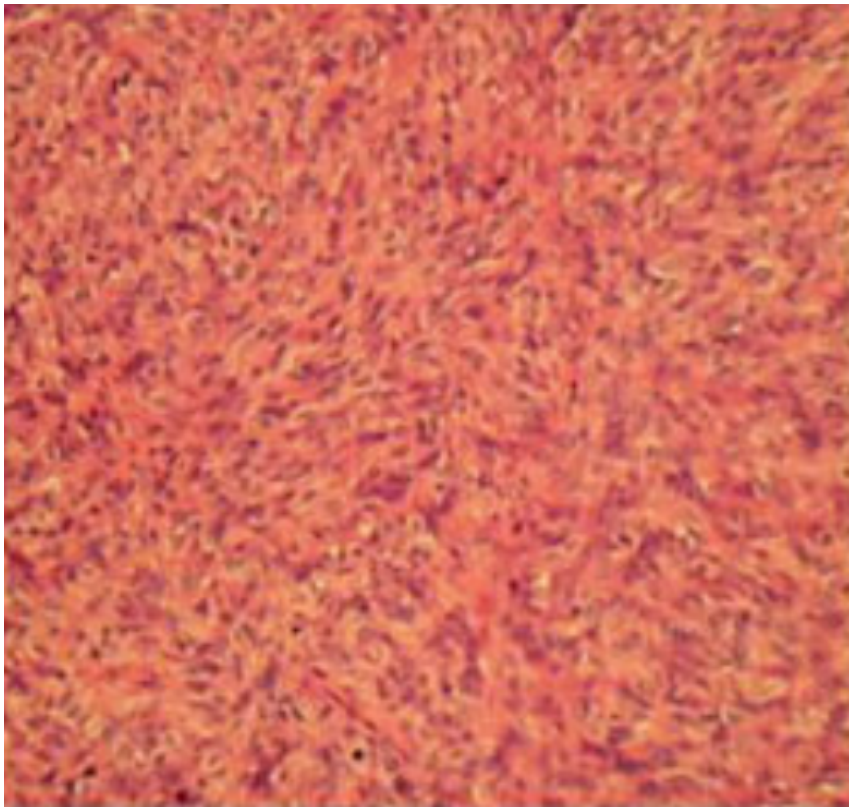



# Malignant transformation in vestibular schwannoma

- Risk $\approx 0.2\%$ of VS cases
- 5-year survival $\approx 20\%$[2]

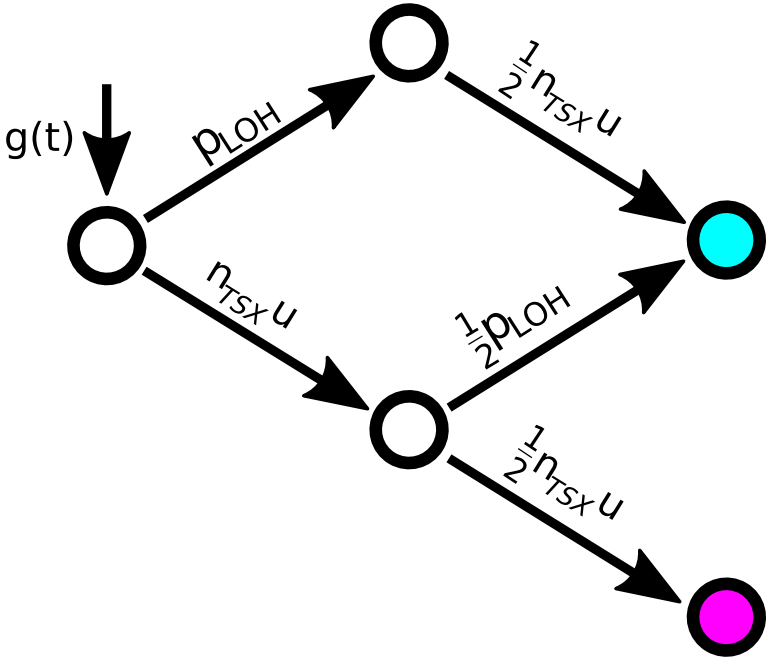


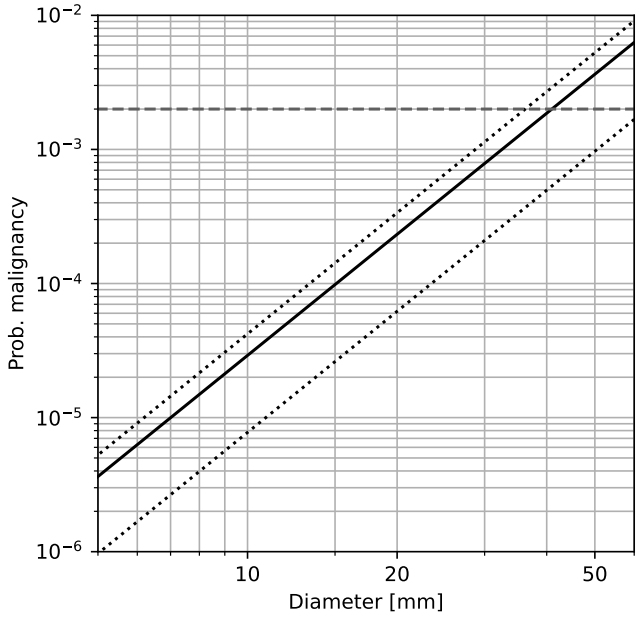

[1] AK Demetriades et al. Skull Base (2010)20:381–387.

[2] R Miao et al. Radiotherapy and Oncology (2019)137:61–70.

# Malignant schwannoma

## Tumour suppressor *TSX* inactivation

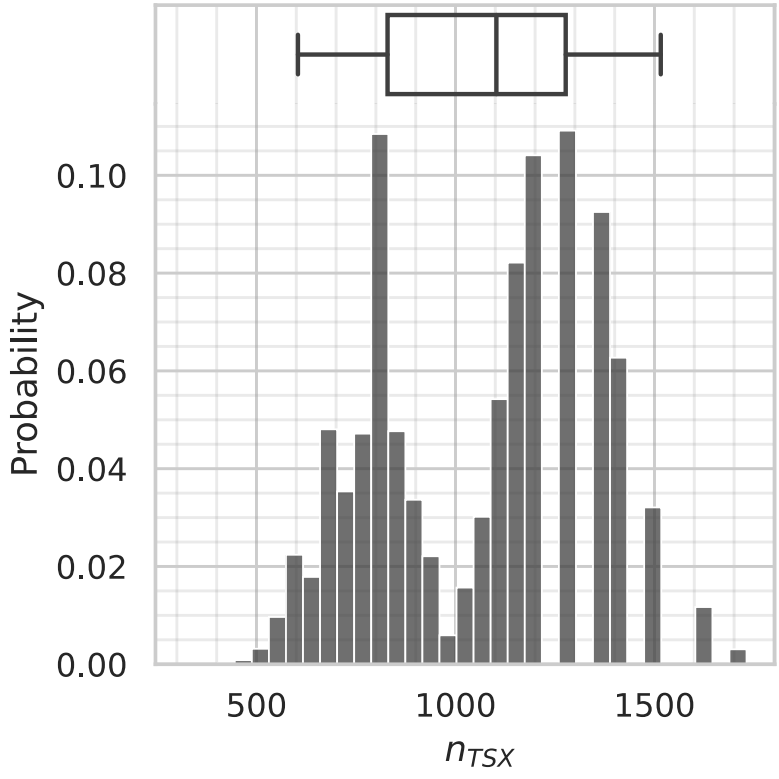



- *TSX* inactivation $\Longrightarrow$ low risk
- Can also estimate $n_{TSX}$ that's consistent with incidence

# Who is *TSX*?

## Parameter estimates for $n_{TSX}$



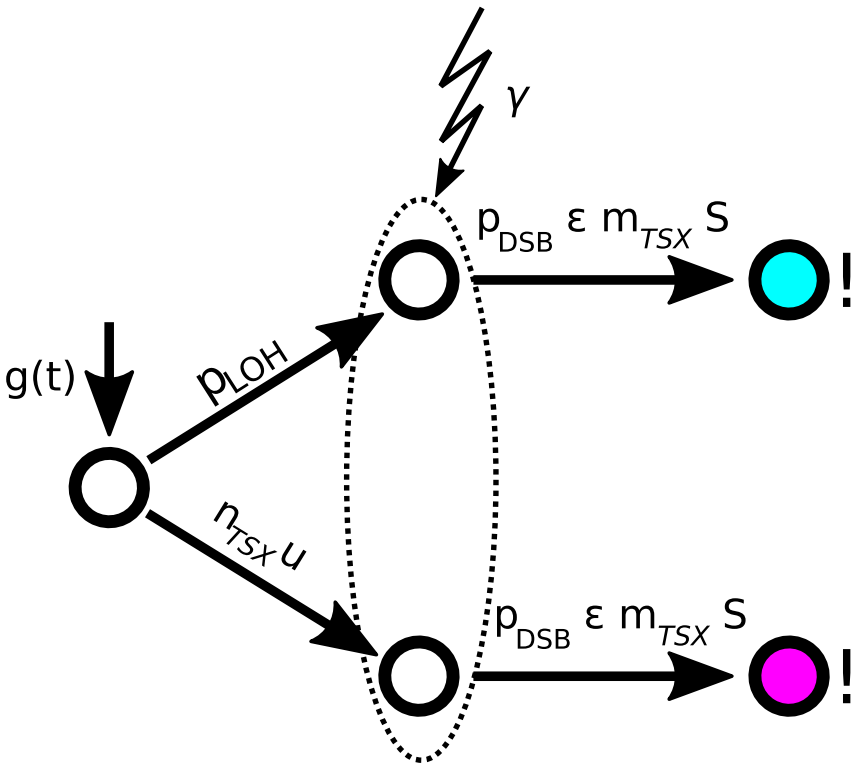

Probably multiple (10?) distinct tumour suppressors

i.e. not (just) *TP53*: $n_{TP53} = 73$

# Malignant schwannoma

## Radiation

Why do we care about *TSX* anyway?



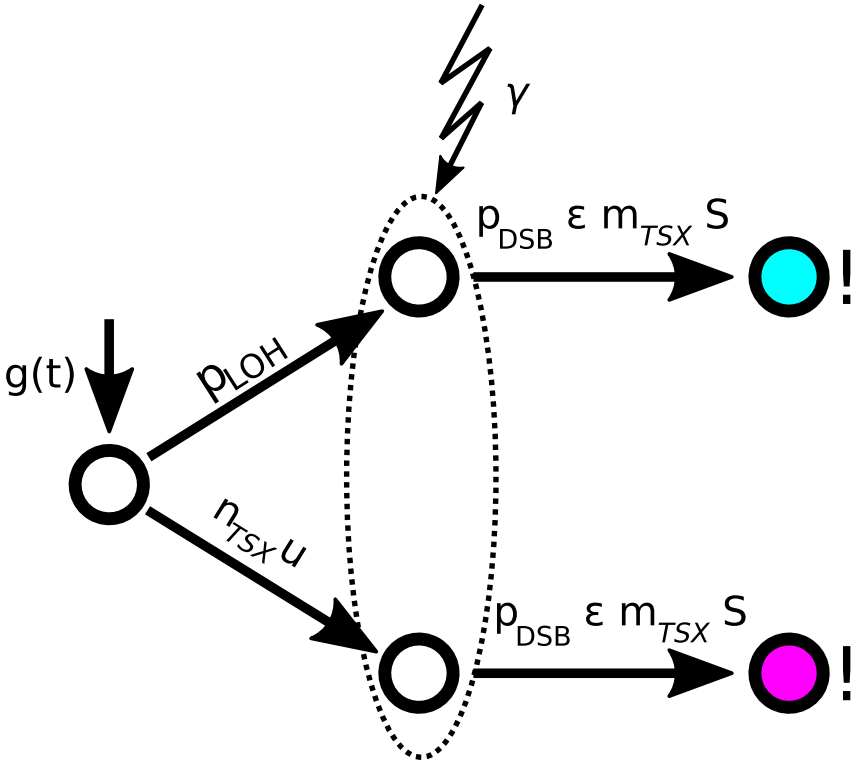

3 dose-dependent effects:

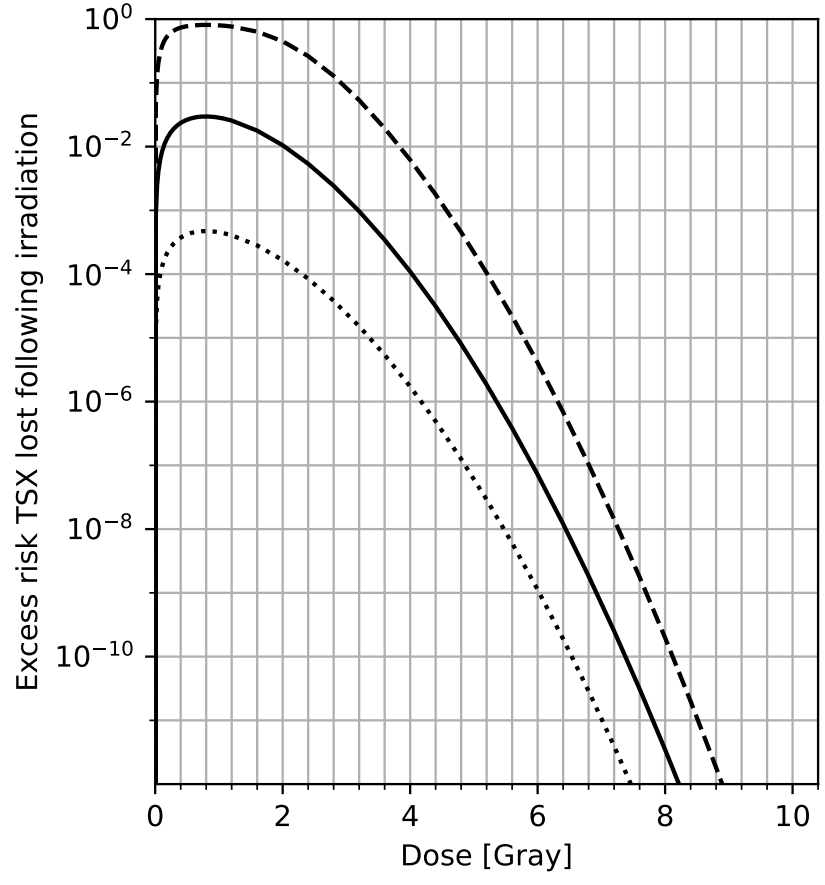
- DSB induction $p_{DSB}(D)$
- DSB misrepair $\varepsilon(D)$
- cell survival $S(D)$

# Malignant schwannoma

## Radiation

Why do we care about *TSX* anyway?



# Main outputs

1. Better estimates of event rates in Schwann cells
2. Can constrain **size** of *GFX* and *TSX* (!)

but...

1. Identity of *TSX* unknown
2. Constraints weak: *GFX* and *TSX* probably multiple genes

# To do list

Next gen sequencing...

- Sporadic VS to constrain $f_{SMARCB1}$: $n > 300$
- CNA/NGS in MPNST (**rare!**): $n > 30$

but also...

- Better models!!! CNA/VAF models, machine learning, optimal clustering, $\mu_{C>T}$ etc.
- SEER data too?
- Multiple genes *GFX* and *TSX*?
- Haploinsufficiency, selection?

Lots to do!

I need collaborators!

# Acknowledgements + collaborators

for their *in kind* support


The University of Washington

In order of appearance...

- Ivana Božić
- Hans Clevers
- Gareth Evans
- Xanthe Hoad
- Miriam Smith