Scaling of scores to ensure gender-independent outcomes

- Women score proportionally lower than men in achieving funding.
- Studies over many years in the USA and Europe have shown that there is a clear gender bias in assessing grants, papers and CVs.
- This bias is independent of the gender of the assessor and has been seen repeatedly in experimental studies over the past 50 years.
- Recent NHMRC results have demonstrated that such a bias also applies to NHMRC funding schemes.

Women do twice as well  

Men do twice as well
Calculating the likelihood that gender-specific funding rates are equal

<table>
<thead>
<tr>
<th></th>
<th>Funded</th>
<th>Not funded</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>1831</td>
<td>7589</td>
<td>9420</td>
</tr>
<tr>
<td>Male</td>
<td>3189</td>
<td>10807</td>
<td>13996</td>
</tr>
<tr>
<td>Total</td>
<td>5020</td>
<td>18396</td>
<td>23416</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>% funded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>19.44% female</td>
</tr>
<tr>
<td>Male</td>
<td>22.79% male</td>
</tr>
</tbody>
</table>

Ratio of female to male funding rates: 0.8531

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square with Yates’ correction</td>
<td></td>
</tr>
<tr>
<td>Chi-square, df</td>
<td>37.27, 1</td>
</tr>
<tr>
<td>z</td>
<td>6.105</td>
</tr>
<tr>
<td>P value</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>P value summary</td>
<td>****</td>
</tr>
<tr>
<td>One- or two-tailed</td>
<td>Two-tailed</td>
</tr>
<tr>
<td>Statistically significant? (alpha&lt;0.05)</td>
<td>Yes</td>
</tr>
<tr>
<td>Strength of association</td>
<td></td>
</tr>
<tr>
<td>Relative Risk</td>
<td>0.8531</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>0.8105 to 0.8979</td>
</tr>
</tbody>
</table>
Funding by Year

Relative rate of funding for females

- 2014 all schemes
- 2013 all schemes
- 2012 all schemes
- 2011 all schemes
- 2010 all schemes
Funding by Type

- Project Grant: 36.9%
- Research Fellowship: 30.7%
- Translating Research: 71.2%
- Career Development: 50.1%
- Early Career (Intl): 54.8%
- Early Career (Aust): 63.9%
Broad Research Area

- Public Health: 59.4%
- Health Services: 63.1%
- Clinical: 38.4%
- Basic Science: 34.9%

Relative rate of funding for females
All schemes – basic vs clinical
All schemes - basic

Male vs Female Funding Rates
All schemes - clinical

Male vs Female Funding Rates

Clinical
Scaling to ensure gender equity

• We suggest that NHMRC funding decisions be adjusted so that the same proportion of male and female applicants are successful.

• This measure, usually termed “scaling”, does not require complex decisions about how to weight different gender-related factors - it is based on the simple proposition that men and women are equally good at medical research and that funding rates should reflect this fact.

• This will not impact on concurrent measures to account for career disruption – but will deal in a practical way with the intrinsic sources of bias that are apparent from the data.
Does scaling interfere with assessment of academic merit?

• There are many well-accepted instances in which academic results are adjusted to account for systematic differences in assessing merit.

• For example, in the NSW higher school certificate, marks in different subjects are scaled to ensure equivalent success rates in different subjects.

• The different marking practices in English vs Maths mean that the raw scores for good students are generally higher in Maths than English.

• But society believes that excellence in both subjects should be rewarded equally, and that scaling is required so that final marks reflect this view.
Suggestions about how to go about scaling

• In the NSW HSC procedure “a linear transformation is applied to the raw HSC marks in each course to set the top mark to a common value. The marks in each course are then standardised to a mean of 25 and standard deviation of 12 on a one-unit basis.”

• There are also added adjustments to deal with subjects with few students.

• Input from professional statisticians will be required to produce a robust procedure for scaling.

• A type of scaling is already in use to “normalise” outcomes from different panels ie the absolute cutoff scores for each panel are adjusted....so we already agree that we should ensure that the same proportion of grants from each discipline are funded...why not from each gender?
The future of scaling

- Hopefully, we will not need to scale in a few years’ time when other measures have made a difference (NIH funding rates are currently gender neutral, a change from the situation >10 years ago).
- The aim is to ensure that the relative funding rate moves to as close to 1.0 as possible in each area NOW, rather than waiting and hoping that things improve. Given that all major schemes have averaged <1.0 for the past 5 years, we have a lot of ground to make up.
- We should not be afraid of the occasional outcome >1.0. We could even choose NOT to scale specific schemes where the ratio is >1.0, since there are so few!