Gender effects in assessment of clinical teaching: Does concordance matter?

Lynfa Stroud
Mahan Kulasegaram
Risa Freeman
Tulin Cil
Shiphra Ginsburg

ICRE
Sept 28, 2019
Background

• Teaching evaluations often sole measure of teaching effectiveness

• Impact on awards, merit pay, promotions

• Lack of relationship between evaluations and actual learning

• Lack of learner expertise to rate teaching

• Vulnerability to bias
Background

- Male physicians rated higher on trustworthiness, competence, professionalism (Ladha, 2017)

- Male physicians rated as more credible providers of formative feedback (Stroud, 2018)

- Gender-based linguistic differences in narrative assessments by students, residents, fellows 2015 to 2016 (Heath, 2019)

- Female faculty rated lower by students across four clinical rotations at one university from 2008 to 2012 (Morgan, 2016)

- Female faculty rated lower by residents across specialties at one university from 2010 to 2014 (Fasiotto, 2018)
Research Questions

1. Does *gender bias exist* in residents’ assessments of faculty teaching in three departments?

2. Is this influenced by *gender concordance or discordance* between the faculty and resident?
Methods

• Resident Assessment of Teaching Effectiveness (RATE)
  • 7-item, 5-point scale
  • Overall used for analysis

• University of Toronto Departments of:
  • Medicine
  • Surgery
  • Family & Community Medicine

• 2016-2017
# Results - Participants

<table>
<thead>
<tr>
<th>Department</th>
<th>Faculty</th>
<th>Female (%)</th>
<th>Individual Ratings</th>
<th>Residents</th>
<th>Female (%)</th>
<th>Gender Discordance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>800</td>
<td>39</td>
<td>5753</td>
<td>552</td>
<td>46</td>
<td>44</td>
</tr>
<tr>
<td>Surgery</td>
<td>377</td>
<td>16</td>
<td>2249</td>
<td>585</td>
<td>38</td>
<td>35</td>
</tr>
<tr>
<td>Family &amp; Community Medicine</td>
<td>672</td>
<td>53</td>
<td>3438</td>
<td>423</td>
<td>62</td>
<td>46</td>
</tr>
</tbody>
</table>
Internal Medicine

4.65(0.84) = 4.64 (0.82)

4.65 vs. 4.62 (p < 0.02)
Surgery

4.28 (0.78) < 4.46 (0.73)

4.26 vs. 4.46 (p < 0.001)
Family & Community Medicine

4.57 (0.72)

<

4.65 (0.65)

4.57 vs. 4.63 (p < 0.0001)

4.56 vs. 4.68 (p < 0.0001)
Discussion

• Teaching ratings very good with ceiling effect

• Complex pattern of results not consistent across departments

• Small differences but potentially big impact
Discussion

• One of may potential biases

• Solutions?
  • Greater exploration of biases
  • Implicit bias training
  • Programmatic assessment of teaching
Acknowledgements

• Dr. Sharon Strauss,
  • Vice-Chair Mentorship, Equity, and Diversity DOM

• Dr. Stuart Murdoch
  • Program Director DFCM

• Dr. Najma Ahmed
  • Vice-Chair Education DOS

• Mr. Edmund Lorens, DOM
• Mr. Haitao Zhang, DFCM
• Ms. Tess Weber, DOS

lynfa.stroud@sunnybrook.ca
References


<table>
<thead>
<tr>
<th>Department</th>
<th>Male Faculty (Mean (SD))</th>
<th>Female Faculty (Mean (SD))</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medicine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male Residents</td>
<td>4.64 (0.82)</td>
<td>4.65 (0.84)</td>
<td>F(1,594)=0.17, p &lt; 0.78</td>
</tr>
<tr>
<td>Female Residents</td>
<td>4.62</td>
<td>4.65</td>
<td>p &lt; 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NSD</td>
</tr>
<tr>
<td><strong>Surgery</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male Residents</td>
<td>4.46 (0.73)</td>
<td>4.28 (0.78)</td>
<td>F(1,313)=13.76, p &lt; 0.001</td>
</tr>
<tr>
<td>Female Residents</td>
<td>4.41</td>
<td>4.48</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NSD</td>
</tr>
<tr>
<td><strong>Family &amp; Community Medicine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male Residents</td>
<td>4.65 (0.65)</td>
<td>4.57 (0.72)</td>
<td>F(1,505)=12.5, p &lt; 0.0001</td>
</tr>
<tr>
<td>Female Residents</td>
<td>4.63</td>
<td>4.57</td>
<td>p &lt; 0.0001</td>
</tr>
</tbody>
</table>
Limitations

• Definition of gender

• 5-point rating scale heavily skewed

• Missing faculty rank for DOS

• One faculty of medicine during one year

• Identified correlations but not causation
Other covariates:
On/Off Service Resident
Faculty Academic Rank
Duration Teaching Exposure

Individual Random Effects
Teacher Gender
Concordance or Discordance
Trainee Gender

Teaching Effectiveness