We recommend that PCa risk prediction models that report discrimination between (all) prostate cancer and no prostate cancer will, almost certainly, also discriminate between clinically significant prostate cancer and no prostate cancer. Indeed this is observed in the few models that have been evaluated [1].

We agree with the authors that the clinical utility of models including prostate volume in prediction models is limited. This is not because prostate volume requires an overly invasive test (transrectal ultrasound, TRUS), but because the test cannot be easily administered in primary care [1] and should therefore be considered as a possible triage of those with equivocal results on their screening test [1]. TRUS itself is not normally associated with significant side-effects or complications as opposed to when TRUS is used to aid prostate biopsy [3], which is well known to be associated with a significant risk of sepsis, bleeding and retention of urine.

K. S. Louie1*, A. Seigneurin1,2,3, P. Cathcart4 & P. Sasieni1

1Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK;
2Université Joseph Fourier Grenoble 1, CNRS, TIMC-IMAG UMR 5525, Grenoble;
3Unité d’évaluation Médicale, Centre Hospitalier Universitaire de Grenoble, Grenoble, France;
4Department of Urology, University College Hospital London & St Bartholomew’s Hospital London & Centre for Experimental Cancer Medicine, Bart’s Cancer Institute, London, UK

*E-mail: k.louie@qmul.ac.uk

accuracy of prostate specific antigen (PSA) screening? [1]. The authors argue that calibration is a more important statistic than discrimination because a patient wants a model that predicts his risk accurately rather than one that distinguishes between men at higher and lower risk. Both aspects are important and models should be evaluated based on both discrimination and calibration. We agree that calibration should be more widely reported, but disagree that it is more important than discrimination. A model that does not discriminate at all may be perfectly calibrated (if it correctly gives the average risk to all men) and such a model cannot be easily improved. In contrast, a model that discriminates well but is poorly calibrated can be fixed by recalibration [2]. We recommend that PCa risk prediction models that report good discrimination should not be used in new populations without out recalibration.

Carlsson et al. imply that we chose only models that predict any PCa for study inclusion in our study. We also reviewed models that predicted clinically significant (aggressive) PCa. However, only the Prostate Cancer Prevention Trial (PCPT) model was externally validated in five or more study populations and met study inclusion criteria (see Table 3 and Figure 4 in [1]). We were also critical in our discussion that despite the overwhelming efforts to develop over 120 unique models, that models have not been built to distinguish between clinically significant (high-grade PCa) and non-significant (low-grade PCa) models to address the problem of overdiagnosis of non-significant tumours. We agree that future developments of prediction models should consider high-grade PCa as an end point rather than any PCa. It should be emphasized that we need models that discriminate between clinically significant and clinically insignificant prostate cancer. Any model that discriminates between (all) prostate cancer and no prostate cancer will, almost certainly, also discriminate between clinically significant prostate cancer and no prostate cancer. Indeed this is observed in the few models that have been evaluated [1].

We agree with the authors that the clinical utility of models including prostate volume in prediction models is limited. This is not because prostate volume requires an overly invasive test (transrectal ultrasound, TRUS), but because the test cannot be easily administered in primary care [1] and should therefore be considered as a possible triage of those with equivocal results on their screening test [1]. TRUS itself is not normally associated with significant side-effects or complications as opposed to when TRUS is used to aid prostate biopsy [3], which is well known to be associated with a significant risk of sepsis, bleeding and retention of urine.

K. S. Louie1*, A. Seigneurin1,2,3, P. Cathcart4 & P. Sasieni1

1Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK;
2Université Joseph Fourier Grenoble 1, CNRS, TIMC-IMAG UMR 5525, Grenoble;
3Unité d’évaluation Médicale, Centre Hospitalier Universitaire de Grenoble, Grenoble, France;
4Department of Urology, University College Hospital London & St Bartholomew’s Hospital London & Centre for Experimental Cancer Medicine, Bart’s Cancer Institute, London, UK

*E-mail: k.louie@qmul.ac.uk

We read with interest the re-analysis by Smith et al. of the randomized, double-blind, phase III trial of zoledronic acid (ZA) versus denosumab [1] and the accompanying commentary [2]. While the former manuscript seeks to reassure physicians that the benefit of denosumab over ZA is still present with the definition of symptomatic skeletal events (SSEs), the latter commentary raises important issues regarding the need to seek more robust measures of the comparative benefits of this class of drugs.

We agree with the commentary, and suggest that ‘radiation to bone’, while indirectly representing a symptomatic event, is a poor objective measurement in the assessment of prostate cancer morbidity. In both, the pivotal trials of antiresorptive therapies, the administration of radiation was not accompanied by any objective assessment of pain to justify radiotherapy: patients could be administered radiation (i) without attempts to control pain with appropriate analgesic medication or (ii) without pain at all to prevent pathological fractures in weight-bearing bones. Additionally, the risk–benefit and cost–benefit ratios of delivering long-term antiresorptive therapy with accompanying risks of hypocalcemia and osteonecrosis of the jaw (ONJ) seem considerably higher than that of short-course radiotherapy to bone. For example, with denosumab, the patient-year adjusted incidence of confirmed ONJ is 1.1% in the first year of treatment, 3.7% in the second year, and 4.6% per year thereafter, and rare but fatal cases of hypocalcemia have been reported [3]. These concerns are especially poignant as there is no robust evidence that either ZA or denosumab are effective in reducing pain. The recognition that only a minority of the patients included in these trials received modern antican-
omitting ‘radiation to bone’ as a component of an SSE. We defined the triad of pathological fracture, spinal cord compression and need for surgery to bone as ‘morbid skeletal events’ (MSEs). With this analysis (Table 1), we demonstrate that while ZA demonstrated a significant delay in MSE compared with placebo ($P = 0.004$, there was no significant benefit of denosumab over ZA for both first and first and subsequent MSE ($P = 0.07$ and 0.35, respectively). Further, risk–benefit analysis (number needed to treat/number needed to harm) [5], suggests that, for example, to prevent one pathologic fracture or spinal cord compression using denosumab, 6–12 additional patients would experience a grade 3/4 toxicity or hypocalcemia when compared with using ZA while 1–2 additional patients would experience ONJ.

Taken together, we suggest that the routine use of antiresorptive agents in men with metastatic castrate refractory prostate cancer, as quantified by SREs and SSEs, is inappropriate if men are responding to anticancer therapy. MSE analysis may be more relevant but we concur with Tombal that more work needs to be done to optimize the choice, timing and length of administration of agents and present our analysis to stimulate further discussion.

R. Leibowitz-Amit1, L. Khoja2, I. F. Tannock2 & A. M. Joshua2*

1Oncology Institute & Cancer Research Center, Sheba Medical Center, Tel-Hashomer, Israel;
2Princess Margaret Cancer Centre, Toronto, Canada
(*E-mail: anthony.joshua@uhn.ca)

disclosure

The authors have declared no conflicts of interest.

references


doi: 10.1093/annonc/mdv039
Published online 28 January 2015

Letter to the editor concerning ‘Trastuzumab emtansine (T-DM1) versus lapatinib plus capecitabine in patients with HER2-positive metastatic breast cancer and central nervous system metastases: a retrospective, exploratory analysis in EMILIA’

It was with great interest that we read the manuscript by Krop et al. [1] published in Annals of Oncology evaluating the efficacy and safety of trastuzumab–emtansine (T-DM1) versus capecitabine–lapatinib (XL) in patients with treated, asymptomatic CNS metastases at baseline and, similar to the EMILIA ITT population [2], T-DM1 was associated with significantly improved OS compared with XL in this subgroup.

Recently, the incidence of brain metastases in HER2-positive breast cancer increased, likely because of the availability of

Table 1. Comparison of MSE across trials

<table>
<thead>
<tr>
<th>$P$ value (Fisher’s exact)</th>
<th>Difference (N)</th>
<th>Placebo (N = 208)</th>
<th>Zoledronic acid 4 mg (N = 214)</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.004</td>
<td>25</td>
<td>67</td>
<td>42</td>
<td>MSE up to 15 months, $n$ (taken from [4])</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>46</td>
<td>28</td>
<td>Pathological fracture</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>Surgery to bone</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>14</td>
<td>9</td>
<td>Spinal cord compression</td>
</tr>
<tr>
<td>$P$ value (Fisher’s exact)</td>
<td>Difference (N)</td>
<td>Denosumab (N = 950)</td>
<td>Zoledronic acid (N = 951)</td>
<td>Event</td>
</tr>
<tr>
<td>0.07</td>
<td>18</td>
<td>38</td>
<td>56</td>
<td>First MSE, $n$ (taken from [4])</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>26</td>
<td>38</td>
<td>Symptomatic spinal cord compression</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>Surgery to bone</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9</td>
<td>13</td>
<td>Symptomatic pathological fracture</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>57</td>
<td>68</td>
<td>First and subsequent MSE, $n$</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>35</td>
<td>43</td>
<td>Symptomatic spinal cord compression</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>Surgery to bone</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>15</td>
<td>17</td>
<td>Symptomatic pathological fracture</td>
</tr>
</tbody>
</table>

MSE, morbid skeletal events.