Are Pathologists Biased? Effect of Blinding on Pathologist Assessment of Lymph-Vascular Invasion in Breast Cancer (Proposal for New Generation Internet-Based Randomized Controlled Trial Design)

Nikita Makretsov*

The University of British Columbia, The University of Oxford, UK

Abstract

Pathology is interpretative diagnostic clinical discipline based on visual evaluation of pattern of tissue alterations in disease as opposed to norm. It is an integrated heuristic process, resulted in formalized pathology report which contains diagnostic and prognostic items, which frequently correlate to each other, and which interpretation is frequently non-independent of each other. Therefore, there is a risk of incorporation bias, which could inflate diagnostic and prognostic value of certain clinical parameters. This could potentially mislead clinicians in their further decision making process regarding appropriateness of chemotherapy therapy for individual patients. Here we propose a novel model for internet-based randomized controlled trial, which would allow to explore the extent of incorporation bias (if any), using pathologists’ interpretation of lymphovascular invasion in breast cancer.

INTRODUCTION AND BACKGROUND

Being a practicing pathologist, I am looking for ways of making pathology diagnostics more reliable and reproducible using Evidence-Based Medicine tools. Randomized trial design remains very uncommon in diagnostic studies, but may resolve the debate whether the pathologists should or should not be blinded of clinical information while providing pathology tests reading.

Lympho-vascular invasion pathology test (LVI) is a fundamental biologic process responsible for a metastatic spread leading to death in cancer patients. The assessment of LVI is done by the pathologists by microscopy of surgically removed cancer tissue. The LVI status of cancer is used by clinicians for risk stratification and planning for systemic cancer therapy, especially when the lymph node status is negative or unknown [1]. Nevertheless, it has been questioned whether LVI data should be used in medical decision making at all due to variation of pathology assessment [2]. The factors which influence pathologist’s decision regarding LVI status determination are not well studied, but behavioral aspects related to pattern recognition may play role [3]. I hypothesize that LVI assessment could be influenced by the knowledge of lymph node status available to pathologists, and could result in incorporation bias in pathology diagnostics, noted earlier [4-6]. The goal of this short essay is to design an RCT with the ultimate goal to detect and measure such bias and to propose the solution to decrease its role (if any) in pathology diagnostics. This task is methodologically challenging, but it has been outlined by [7,8].

PICO

Is knowledge of lymph node status in breast cancer influences the pathologist evaluation of LVI in breast cancer?

Population (participants)

Inclusion criteria: pathologists (representative sample of practicing Canadian pathologists stratified by age, sex, years
in practice, rural or urban settings, subspecialty of general pathology practice).

**Exclusion criteria:** non-practicing pathologists, pathology trainees and biomedical researchers, pathologists whose practice does not include diagnosis of breast cancer.

**Intervention**

the sets of digitalized de-identified to pathologists slides of breast cancers (cases) with blinded information on lymph nodes. There will be whole standard sections hematoxylin-eosin stained slides, scanned by whole-slide digital image scanner (scanning magnification x20) Aperio, Hitachi or any other equivalent model. The quality of scanned images will be assessed by an independent qualified pathologist, who will be kept blinded of the goals of the study. The scanned images should be equivalent or superior in quality compared to standard light microscopy, as variation attributed to optical artefacts and aberration will be absent.

**Comparison**

The same source set of cancer slides with unmasked lymph nodes.

The participants will be asked to review the cases the way they do in practice and provide the assessment of the given slides and determined LVI along with 2-3 other high risk parameters (distractors).

**Outcome**

Binary: positive or negative LVI status, using current [1,9] guidelines.

**The gold standard**

For LVI status could be additionally determined by expert consensus, but is not essential and is out of scope in this trial.

**Trial Design**

Standard “p x q” cross-over design [10,11], where p- sequence of intervention(A)-control(B) cases administered, and q- repeat exposures of pathologists to different breast cancer cases (slides), summarized in Table 1.

**Allocation concealment**

Recruitment of participants will be shielded from the nature and sequence of cross-over periods.

**Blinding**

The identity of all the pathology slides will be masked to the participants. Unique bar-codes will be generated for each “q” cross-over and participant, to ensure adequate blinding. True slide identity codes will only be available to central registry. In addition, the assessors of the results (statisticians) will be blinded of the nature of the cross-over periods, while performing statistical analysis.

**Randomization**

Central internet-based randomization technic will be utilized, using the database of digital images. The automated algorithm will chose cases at random and generate unique cross-over sequence of exposure (A)/controls (B) for each participant and batch-period.

**Definition of cross-over periods**

- Period A (Intervention): Blinded assessment - only tumor slides will be given, no lymph node slides will be provided.
- Period B (Comparison): Un-blinded (traditional) assessment of LVI in breast cancer: tumor slides will be given along with lymph node slides from the same patient.

**Sample size determination**

There are few challenges I face while attempting power calculations, due to my limited statistical expertise:

Firstly, there should be a valid number of pathologists (participants).

Secondly, each pathologist should be exposed to a valid number of slides with and without LVI (LVI positive case prevalence in real practice is 15-25%). This could make the trial large and impractical, thus LVI-enriched slide set with close to 50% LVI-positive prevalence might be necessary.

**Regular parallel 2-arm trial sample size calculation:**

Assuming that blinded LVI-positive assessment (intervention arm) event rate is 20%, and in un-blinded assessment (comparison arm) it is at least 40%, the difference in the event rate could be as much as 20%.

---

**Table 1:** Trial Design Summary.

<table>
<thead>
<tr>
<th></th>
<th>Control – not blinded (B)</th>
<th>Intervention – blinded (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event (LVI-positive)</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>No event (LVI-negative)</td>
<td>c</td>
<td>d</td>
</tr>
</tbody>
</table>

Control event rate (CER)=a/(a+c) Experimental event rate (EER)= b/(b+d)
In order to detect 20% difference, 172 participants (pathologists) are required to have a 80% chance of detecting, as significant at the 5% level, an increase in the primary outcome measure from 55% in the control group to 75% in the experimental group, if there would be a parallel design RCT. The online power calculator was used (www.sealedenvelope.com/power).[12]

It will make it difficult to recruit such large cohort of pathologists and will make the study cost-prohibitive.

The alternative 2-period crossover design would require $172/2 = 86$ pathologists. Considering correlation between the repeat measures (as all the cases will be given twice, under different cross-over periods), the sample size could be further reduced, assuming the correlation coefficient in the range of 0.3-0.5.

Thus, the final sample size using the formula $N(\text{crossover}) = (1-r)^2N(\text{parallel})/2$ will be in between: $(1-0.5)^2*172/2 = 43$ and $(1-0.3)^2*172/2 = 60$. Allowing for 10-15% attrition, the study will require 43+ (43/100)15 = 49 to 60+ (60/100)15 = 60+9 = 69 pathologists-participants. The number of pathology slides will have to be determined by a qualified statistician. The McNemar’s test for a paired dichotomous response could be used for such task [10,11].

**Trial Management**

The trial will be delivered through the internet. Rigorous evaluation and testing of randomization algorithm will be required prior to opening of the trial to the participants, to ensure the correct linkage between the true identities of the slides, cross-overs and the individual participant results.

The histologic slides will be digitalized and entered into a central database. This will avoid shipping of perishable glass slides and will ensure blinding, randomization of cross-over, internet-based delivery simultaneously to multiple participants and instant data collection.

Each participant will receive a unique password and will be provided with a unique digital slide set (see below), as per the randomization algorithm.

The trial will be divided into several runs (batches), as determined by number of parallel repeat observations based on sample size required.

The assessments will be due within a week, for a reason of practicality.

In order to preserve blinding within each batch, the repeats of the same cases with different cross-over will be avoided by using a computer algorithm. Each subsequent batch will be divided by a wash-out period of 4 weeks.

**Monitoring of process**

The participants will be notified electronically when the batches of digital slides are ready for assessment. The results will be submitted online and instantly collected by central registry. In case the results are overdue, the reminder will be sent, and if no reply is received within the specified time frame, the phone call will be delivered. The reasonable extension for submissions could be granted.

**Analysis**

We will need to determine whether the relative risks of making a positive LVI diagnosis is dependent on the knowledge of lymph node status of the patients. The standard $2 \times 2$ table (Table 2) will be populated and the results will be analyzed using the formulas:

- Relative risk (RR): CER/EER
- Absolute risk reduction (risk difference): CER- EER
- Relative risk reduction: (CER-EER)/CER

95% confidence intervals will be calculated for each parameter as described by [13,14].

The RR (95% CI) including 1.0 will signify the absence of incorporation bias in pathology assessment of LVI, within the pre-specified alfa and power.

**Reporting**

Trial results should be published in an open source medical venue.

**SUMMARY**

Pathology is interpretative diagnostic clinical discipline based on visual evaluation of pattern of tissue alterations in disease as opposed to norm. It is an integrated heuristic process, resulted in formalized pathology report which contains diagnostic and prognostic items, which frequently correlate to each other, and which interpretation is frequently non-independent of each other. Therefore, there is a risk of incorporation bias, which could inflate diagnostic and prognostic value of certain clinical parameters. This could potentially mislead clinicians in their further decision making process regarding appropriateness of chemotherapy therapy for individual patients. Here we propose a novel model for internet-based randomized controlled [15] trial, which would allow to explore the extent of incorporation bias (if any), using pathologists’ [16] interpretation of lymphovascular invasion in breast cancer.

**REFERENCES**

5. Gunia S, Berg T, Gradhed E, Becker S. Knowledge of the anatomical polypl location might bias the pathological classification of histologically equivocal colorectal serrated polyps - a consensus study
performed by pathology trainees. Pathology Research and Practice. 2011; 207: 116-120.


