Desiderata for Major Eligibility Criteria in Breast Cancer Clinical Trials

Matthew L. Paulson, MPH, Chunhua Weng, PhD
Department of Biomedical Informatics, Columbia University, New York, NY

Abstract
Use of major eligibility criteria is a popular but unstudied folk practice for improving patient screening efficiency for clinical studies. This mixed-methods research study derived the desiderata for major eligibility criteria in breast cancer clinical trials. We randomly selected thirty interventional breast cancer clinical trials conducted at The New York-Presbyterian Hospital on the Columbia University Medical Center campus to create training (N=20) and testing (N=10) datasets. We utilized the Think-aloud protocol to gauge how clinical researchers identify and use major eligibility criteria to prescreen patients for clinical trials during an audio-recorded interview. A focus group session was held to understand the current prescreening process and investigate how it could be optimized to maximize recruitment rates. Using the grounded theory method, we annotated transcriptions to discover user rationale and desiderata behind major eligibility criteria in breast cancer clinical trials, which were evaluated later in a follow-up survey.

Introduction
Recruitment to clinical trials remains the biggest barrier to clinical and translational research. Most clinical trial studies have dozens of complex inclusion and exclusion eligibility criteria. Therefore, when using these eligibility criteria for subject screening, a common cost-effective practice adopted by clinical researchers is to start with a small set of “major eligibility criteria”. If a patient does not satisfy any major criterion, there will be no need to waste the time to go over the complete eligibility criteria list. Moreover, many clinical research coordinators know that a lot of minor eligibility criteria are either irrelevant to or rare among the target population so that the information gained from those minor criteria returns minimal value for screening eligible patients. Therefore, they focus on using major eligibility criteria to maximize information gain during prescreening processes.

Understanding the definition and characteristics of major eligibility criteria has three foreseeable benefits: (1) facilitating automated selection of major eligibility criteria for efficient electronic clinical trial prescreening using electronic patient data; (2) enabling cost-effective analysis of existing eligibility criteria; and (3) enabling understanding of the prioritization behaviors of clinical researchers to inform the design of advanced clinical trial screening methods. However, to the best of our knowledge, little is known about the selection of major eligibility criteria for clinical trial prescreening. With a focus on the important disease area of cancer, in this study our research questions are “is there a common definition for major cancer clinical trial eligibility criteria?” and “if yes, what are the characteristics of these major eligibility criteria?” In this study we define a desiderata for major eligibility criteria of clinical trials as well as to provide examples of standardized major eligibility criteria using a mixed-methods approach using qualitative and quantitative measures with clinical researchers at Columbia University Medical Center Cancer Center. The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects, and was approved by the Columbia University Medical Center Institutional Review Board.

Background
1. Barriers to recruitment and the need for prescreening

Slow clinical trial recruitment is a significant obstacle to medical discovery. Recruitment obstacles can lead to delays from one to six months for most clinical trials. Automating a significant portion of the patient screening process could reduce recruitment delays by identifying potentially eligible patients. The lack of standardization for eligibility criteria creates obstacles to comparing trials to one another. As a result, patient screening is a complex multi-step and multi-user process that requires significant and expensive human time and effort. Moreover, the complexity of eligibility criteria can cause potentially eligible patients to be overlooked. Information systems have the potential to decrease delays in clinical trial recruitment and minimize human resources that manually review records for prescreening. Weng et al. showed that a real-time screening alert before a clinical encounter for patients improved the rate of recruitment over a twelve-month period by identifying patients that are “potentially eligible” for full screening for a clinical trial. Prescreening of patients for clinical trials requires use of study specific predefined inclusion and exclusion criteria.

2. Related work
Most prior efforts have focused on the standardization of eligibility criteria with expressive ontologies. For example, the Eligibility Rule Grammar Ontology (ERGO) provides a formal representation for eligibility criteria. ERGO is not sensitive to the complexity of criteria and does not represent the stability of eligibility data elements or support fuzzy representations such as “potentially eligible”. A related effort called the eligibility criteria extraction and representation (EliXR) extended ERGO by providing richer semantic information and decomposing complex eligibility criteria into more meaningful semantic segments. EliXR-Time further extended EliXR by supporting temporal representations. For example, a temporal constraint might require a medication washout period of three weeks. Such efforts have focused on rich semantics and full expressiveness of the criteria, though none of them has been widely adopted yet or have generated real-world impact on clinical trial recruitment processes. In this study, we took a novel angle to this problem by addressing a real user need for classifying major eligibility criteria, which has been observed in folk practice of clinical trial prescreening but has remained largely unstudied.

3. Challenges of cancer clinical trial recruitment

Unique challenges exist for cancer clinical trials. The knowledge of cancer is increasing at an unprecedented rate along with its understood complexity. Patients that can be recruited for cancer clinical trials vary widely in terms of comorbidities, overall health, and cancer pathology. Previous studies have indicated that eligibility criteria concepts may be disease specific, and differences in patterns of concepts for eligibility criteria were found between breast cancer and cancer in general. As a result, eligible patients must meet complex criteria, which are not standardized and vary by disease. In addition, not all cancer patients react similarly to the same therapeutic treatment. Heterogeneity in patient response to cancer therapies likely is caused by inter-individual differences in drug deposition and pharmacokinetics. As we enter the era of precision medicine, patients will not be assigned to drug trials without attempting to predict how they will respond to the therapeutic regimen. This complex problem will require efficient information systems. Informatics solutions will need to be built within the clinic workflow. Major eligibility criteria offer one effective strategy for streamlining clinical trial recruitment in cancer centers and to potentially increase patient satisfaction by limiting wait-times and maximizing care quality by quickly presenting all available treatment options to patients during clinical encounters.

4. Our previous pilot study

We previously conducted a pilot study involving nine clinical researchers, including physicians, nurses, and clinical research coordinators. Participants were invited to take a survey about how they use eligibility criteria to screen patients for clinical trials. Our study revealed a concept known as “major eligibility criteria”, which could be useful to assist investigators by automatically prescreening patients with data kept in the electronic health record. However, the responses revealed no clear consensus for the definition of “major eligibility criteria”. Each expert panelist defined the concept slightly differently, yet we identified frequently cited major criteria, such as histology and disease stage of cancer, in the aggregate data. Trends also emerged regarding how the panelists used major eligibility criteria in practice. For example, panelists described how they use major eligibility criteria to prescreen patients. The results also uncovered disparities about which criteria should be considered major. Our pilot study helped us to identify the unanswered research question and informed this study, which is to perform in-depth interviews and a focus group to understand the design requirements for prescreening with major eligibility criteria. The pilot study also suggests that major eligibility criteria can potentially improve triage efficiency for clinical trial recruitment by identifying patients that are potentially eligible for clinical trials. However, to the best of our knowledge, no prior study has investigated the major eligibility criteria for clinical research recruitment and discussed user requirements for this type of decision support. This study employs a mixed-methods approach that includes both qualitative and quantitative measures to understand how clinical researchers define, select, and use major eligibility criteria. Moreover, we harness their collective knowledge to inform future informatics designs for triaging patients for clinical research eligibility prescreening.

Methods

1. Study setting

We recruited physicians, nurses, and study coordinators from the Breast Cancer Program in the Herbert Irving Comprehensive Cancer Center (HICCC) at Columbia University Medical Center (CUMC). CUMC is an academic medical center in New York City that is affiliated with the New York-Presbyterian Hospital. Clinical researchers in the Breast Cancer Program at HICCC were sent an IRB-approved recruitment email with details about the study and invited to request more information. The protocol activities consisted of one initial interview that was audio recorded, one focus group session, and one survey after the focus group session that study participants completed independently. The invitees were informed that completing the survey was voluntarily and they could opt out if they did not wish to participate. Eleven people requested information and were subsequently consented to the research
study. We recruited eleven clinical research professionals, including five physicians, two nurse practitioners, one nurse, and three research coordinators. Two participants were male and nine were female. Their work experience with breast cancer clinical trials was an average of 6.8 years with a median of 6 years. All participants had experience in recruiting patients for breast cancer clinical trials. Figure 1 graphically depicts our study schema.

2. The eligibility criteria dataset

Thirty clinical trials were randomly selected from the portfolio of active interventional breast cancer clinical trials, which represented over 41% of the 72 available cancer trials at the point of study at CUMC. Our random sample included a mixture of protocol phases to create a representative sample. All of our selected trials purposely involve interventions done to the patient. Twenty trials were designated as the training dataset, while the remaining ten were the testing dataset. Since there exist significant differences in the content and word counts between full text protocols and their summaries on CT.gov, which casts doubt on the reliability of clinical trial summaries on CT.gov, we decided to use eligibility criteria from full-text clinical trial protocol documents for this study.

3. The initial interviews

Ten research members participated in the initial interview to review eligibility criteria for six clinical trials and classified each criterion as either major or minor, so that each trial has three raters. The response rate was 100%. The survey was conducted using Qualtrics, an online survey platform (www.qualtrics.com). Participants were encouraged to “think aloud” while they completed the task so we could capture their thought process with audio recording, which was later transcribed. All identifying information was stripped from the audio recordings to ensure anonymity. The answers were downloaded from the Qualtrics website into Excel for analysis. The answers for the major eligibility criteria were analyzed individually for themes and then annotated by author MP, who is a subject matter expert in breast cancer clinical trial data collection with over 9 years of experience. The initial interview completion rate was 100%. We assigned raters using a cascading structure to maximize variability in rater assignments, which ensured that each rater works with different people to maximize result reliability.

4. The focus group

The focus group was held one week after completing the initial interviews. Eight participants, including five physicians, a nurse practitioner, a nurse, and a study coordinator, participated in the focus group. The focus group method was used to find areas of agreement and disagreement and participants were encouraged to have a discussion among themselves to complete the task. The participation rate of the focus group was 100%. The group started the session by drawing their perception of how the current pre-screening process works. The members worked individually and were told to focus on how physicians, nurses, and study coordinators interact with one another during the prescreening process, if at all. Participants were also asked how, when, and where prescreening occurs.

Next, the group reviewed eligibility criteria that were considered major from the initial interviews using two steps. First, the results of selected trials were reviewed with all focus group participants at the same time. During the discussion, Author MP listed the qualities and attributes of major eligibility criteria to learn why participants considered some eligibility criteria to be major. Author MP updated the list as the conversation progressed during the focus group. Second, the focus group reviewed their discussion and agreed upon a list of attributes for major eligibility criteria based upon their subject matter expertise and the discussion of the focus group. This data was recorded using an easel pad that was visible throughout the focus group session, where anyone could write a quality or attribute about major eligibility criteria during the conversation. Then, the focus group reviewed the attributes and came to a consensus on a final list of attributes. Finally, the focus group participants discussed and drew their
version of an optimized prescreening scenario as a group using easel pads that were hung on the wall, so all focus group members could see the diagram and participate.

5. The follow-up survey

The follow-up survey was designed to evaluate the standardized major eligibility criteria and desiderata that were derived from our content analysis of the initial interviews and focus group session. First, we extracted all eligibility criteria from the training data set of twenty breast cancer clinical trials that were considered major by all three raters and referred to them as consensus major eligibility criteria. Second, we reviewed the list of attributes created by the focus group for major eligibility criteria. To ensure saturation of major eligibility criteria attributes, we reviewed and annotated the transcripts of the initial interviews by open adjudication performed by both authors until saturation was reached. Both annotators reviewed the transcripts separately and discussed their results before coming to a consensus on major eligibility criteria attributes. The result was the desiderata for major eligibility criteria in breast cancer clinical trials.

Ten participants completed the follow-up survey two weeks after the focus group session. The follow-up survey was administered using the Qualtrics survey tool and consisted of three parts. First, study participants were asked to classify each standardized major eligibility criterion that was created from the initial interviews using a 5-point Likert scale with options “Definitely Major”, “Possibly Major-Depending on the Context”, “Undecided”, “Possibly Minor”, and “Definitely Minor”\[6\]. The second part of the follow-up survey asked study participants to match examples of standardized eligibility criteria with each attribute of our derived major eligibility criteria desiderata. From the training dataset, twelve major eligibility criteria appeared in the training dataset more than once, so we included them in follow-up survey to have study participants match them with the derived desiderata. Study participants were directed to select as many or as few examples of major eligibility criteria for each attribute in our desiderata. The final part of the follow-up survey asked study participants to classify eligibility criteria as major or minor. Ten breast cancer clinical trials were randomly assigned to our ten participants so each participant would classify the eligibility criteria as major or minor for three testing clinical trials. We annotated the major eligibility criteria in the follow-up survey to see if any additional standardized major eligibility criteria were detected.

6. Data collection and analysis

Study data was collected with three methods: an online survey tool, Qualtrics, where the results were downloaded into Excel; audio recording and subsequent transcription; and the focus group session that created individual and group prescreening workflow artifacts and easel pad drawings. We utilized the “Think-aloud” Protocol to record the thought process of the study participants while they classified the training set eligibility criteria as either major or minor in the initial interview[6]. Then, we analyzed and annotated the audio transcriptions for the initial interviews using grounded theory to ensure saturation of the desiderata for major eligibility criteria from the focus group easel pads[6]. Finally, we asked study participants to evaluate the standardized major eligibility criteria as well as the desiderata in the follow-up survey.

Results

1. Qualitative study results

The average, median, maximum, and minimum time spent on the interview per participant was 22 minutes 42 seconds, 19 minutes 48 seconds, 40 minutes and 1 seconds, and 9 minutes and 42 seconds, respectively. Our interview transcript annotation revealed that all study participants differentiated major and minor eligibility criteria. There was 100% agreement amongst the participants that not all eligibility criteria should be given equal weight in the prescreening process. Participant 07 stated, “In my mind major eligibility criteria, I mean like who’s actually worth approaching upfront”. All study participants were able to find one or more major eligibility criteria for every clinical trial that they rated. Several major eligibility criteria were noteworthy and were reported below.

Disease staging was considered a major eligibility criterion for 100% of the study participants. Disease staging appears to be the most important factor to divide patients into subgroups for treatment decisions and recruitment for clinical trials. Participant 03 stated, “So, the disease staging, I would consider major. That’s your number one.” Participant 05 stated, “staging, this is probably one of the most major, most important”. This finding was confirmed by both transcription annotation and the follow-up survey, where 100% of participants identified disease stage as “Definitely Major”.

Transcript annotation also revealed a contextual sentiment about age. Several study participants stated that age could be considered a major or minor eligibility criterion depending on the context, such as how the numerical value of the age creates subgroups in the overall breast cancer population. For example, the age criterion can be major when its threshold is meaningful, such as “greater than 50 years old”. In contrast, the age criterion “adults greater than 18
years old” is not major because the majority of the patient population satisfies this criterion. Participant 04 stated, “I’m assuming that there’s going to be (over 18)”. Participant 09 stated, “…most of the patients that we see are over 18, so that’s usually an assumption I make that I would call minor, definitely”. The reasoning for this is the degree to which the age value divides the overall breast cancer population into meaningful subgroups. Participant 03 stated, “Most people at least in this setting are going to be 18. When the cut off is larger, like 50, it’s when I would consider it more of a major (criterion)”. 

Physician, nurse, and study coordinator participants did not consider rare phenomena to be major eligibility criteria. Regarding previous cancers, participant 06 stated, “…it wouldn’t be considered a major criterion because for most patients, that’s not a major issue”. Participant 01 speaking about Hepatitis and HIV stated, “While that’s important, it’s not common”. Participant 03 stated, “Bilateral malignancy is relatively uncommon, I would consider that minor”. Participant 09 stated, “I would call that minor because I would say that the majority of the population is not HIV-positive”. Participant 02 stated, “The things that are less common become in my mind not major”.

An area of contention in the transcript annotation was laboratory results. Specifically, a difference existed between a study coordinator and nurses. Participant 08, a study coordinator, stated, “I can assess the labs…so that’s a major to me”. Participant 09, a nurse practitioner, stated, “I would say in terms of pre-screening, labs are minor because the lab is usually the one we use when they actually come in for their screening visit”. Participant 07, a nurse practitioner, stated, “let’s say the patient is essentially eligible except for some lab variations, for the most part in my experience, this can be remedied”. However, the importance of laboratory results may increase for patients with advanced disease. Participant 07 stated “I feel as though the laboratory parameters…have more implications for patients who are treated in a metastatic setting because those women tend to be more sick”.

A major eligibility criterion that generated significant conversation in the focus group session was around the concept of “prior treatment”. While prior treatment was not always considered to be accessible in the electronic health record, the focus group did consider it a major eligibility criterion that was available. However, it might take additional time to receive records from other clinicians that treated the patient. Focus group participants clarified that a patient’s prior treatment can be difficult to find, especially if they have been treated at multiple institutions. The focus group stated that the importance of prior treatment differs between early stage and advanced stage breast cancer patients and hence whether prior treatment is major is also a contextual decision.

The disease staging of the patient may increase the likelihood of other eligibility criteria being considered major. For example, the focus group agreed that the importance of prior treatment proportionally increases in connection with the number of patient prior treatments. For a patient with multiple prior treatments, it is likely that prior treatment becomes a major eligibility criterion for that specific patient. Similarly, if a patient has not had any prior treatments, the concept of prior treatment is less important for that individual patient. This sentiment was also reflected in the transcription annotation. Participant 02 stated during the initial interview, “Given that this is a neoadjuvant study, most patients would not have had prior therapy. So probably not so important to know that they haven’t had prior treatment”. Therefore, prior treatment is more relevant to patients with advanced disease than for early stage disease.

2. Desiderata for major eligibility criteria

Desiderata for major eligibility criteria (Table 1) were derived based upon the transcript annotation from the interviews using grounded theory as well as from a group activity that was conducted during the focus group session. The desiderata were then evaluated in the follow-up survey. The desiderata attributes are ranked in descending order of frequency determined by counting the matched eligibility criteria for each term in the desiderata for all participants that completed the follow-up survey. The twelve most frequent major eligibility criteria from the training data set were used for the major eligibility criteria examples. In addition, the major eligibility criterion with the highest score from study participants is identified as a “Best Criterion Example(s)”, where we have included more than one criterion in the event of a tie.

3. The prescreening process

During the initial interviews, participants identified different phases of patient screening. Participant 06 stated that major eligibility criteria are used for prescreening, while minor criteria will be used during manual reviews of data to give higher scrutiny to the patient. An example of a minor criterion is asking a patient if they would consider participating in a research trial. The transcript annotation revealed that different phases of patient screening occur at different time points. Participant 06 identified three different steps to screening. The first step is looking at “high level variables”. The second step is when you give a patient a more in-depth look for obvious problems that would prevent them from enrolling in the clinical trial such as organ failure. Thirdly, you have the official eligibility checklist, where you go line by line and see if they meet all eligibility criteria.
At the beginning of the focus group session, study participants were asked to draw the workflow of how patients are currently prescreened in their clinics. 100% of the focus group participants completed the task. Seven out of eight participants identified the first step of the prescreening process starting with the physician, where the physician directs nurses or study coordinators to do a manual screening of the patient. The one remaining participant stated that the prescreening could start with any clinical researcher, depending on the type of study. For example, study coordinators may recruit participants for behavioral or observational studies without physician input.

Table 1. Desiderata of major eligibility criteria for breast cancer clinical trials in descending order of frequency.

<table>
<thead>
<tr>
<th>Desiderata</th>
<th>Operational Definition</th>
<th>Rationale Quotes from Study Participants</th>
<th>Best Criterion Example in the Follow-up Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reliable</td>
<td>A good biomarker</td>
<td>“I guess the question is how accurate are the data that are going to be in the EMR.”</td>
<td>Disease staging, HER2 status; Hormone receptor status</td>
</tr>
<tr>
<td>Objective</td>
<td>Assessed in a standardized way without opinions</td>
<td>“one way for me to delineate majors versus minors, ...something that’s like blatantly objective, it’s major.”</td>
<td>Age</td>
</tr>
<tr>
<td>Relevant to me</td>
<td>Something I can review on my own</td>
<td>“that’s going to be pathology review, and I don’t have experience on that”</td>
<td>Eligible to receive treatment</td>
</tr>
<tr>
<td>Available</td>
<td>Information ready to use</td>
<td>“there’s always things that aren’t going to be in the medical record or things that you’re not going to know.”</td>
<td>Disease staging</td>
</tr>
<tr>
<td>Stable</td>
<td>No fluctuation in the short term</td>
<td>“If there is something that the patients can discontinue prior to randomization, then that would be minor.”</td>
<td>Diagnosis made by core biopsy; Gender</td>
</tr>
<tr>
<td>Intrinsic patient trait</td>
<td>Something that cannot be changed by external force</td>
<td>“let’s say the patient is essentially eligible except for some lab variations, for the most part in my experience, this can be remedied.”</td>
<td>Age; Gender</td>
</tr>
<tr>
<td>Accessible</td>
<td>Understandable with little effort</td>
<td>“things that would be easy to prescreen”</td>
<td>Age</td>
</tr>
<tr>
<td>Relevant to now</td>
<td>Something I can find out now</td>
<td>“that won’t happen until we actually meet them.”</td>
<td>Age; Gender</td>
</tr>
<tr>
<td>Prevalent in population</td>
<td>Concepts common to breast cancer patients</td>
<td>“wouldn’t be considered a major criterion because for most [breast cancer] patients that’s not a major issue”</td>
<td>Disease staging</td>
</tr>
<tr>
<td>Must have</td>
<td>Not optional</td>
<td>“Must be post-menopausal, major.”</td>
<td>Disease staging</td>
</tr>
<tr>
<td>Differential for population subgroups</td>
<td>Able to divide population into subgroups</td>
<td>“…age &gt; 18 is not a major criterion but specific age constraints such as between 50 and 75 are major criteria”</td>
<td>Disease staging</td>
</tr>
<tr>
<td>Plausible (to happen)</td>
<td>Something realistic and practical</td>
<td>“We would never approach a person with life threatening metastasis.”</td>
<td>Eligible to receive treatment; Scheduled to receive treatment</td>
</tr>
</tbody>
</table>

The final group activity of the focus group was to describe an optimized prescreening workflow. The prescreening process utilizes major eligibility criteria to automatically identify potentially eligible patients in the electronic health record. Study participants came to a consensus that decision support for clinical trials must be built into the everyday workflow of the clinicians. At CUMC, the clinicians view a patient schedule that is built for each physician’s clinic day. The focus group agreed that research staff should then be able to easily see potentially eligible patients in the physician’s daily schedule. This enables research staff to screen patients and communicate with one another to make an ultimate screening decision.
Figure 2 shows the optimized prescreening workflow as developed by consensus during the focus group session. In addition to describing the workflow, focus group participants also illustrated the type of information that would be useful in the user interface (Yes; No; Maybe). The relevant clinical trials for each participant are listed on the physician’s schedule. For example, Patient 1 is scheduled to see the physician at 8:00 AM and is potentially eligible for “Trial A” and “Trial B” as seen in Figure 2. The automated prescreening process, study coordinator, research nurse, and physician can each select “No–Not Eligible, “Yes–Possibly Eligible”, and “Maybe Eligible” after their respective reviews.

4. Quantitative study results

The quantitative results include the instance counts from the training and testing datasets as well as the Likert scale measurements from the follow-up survey. For the training dataset that composed twenty breast cancer trials, the average count of eligibility criteria per trial was 19.7 with a median of 18. There were more than double the amount that were classified as minor than were classified major. The average major criterion per clinical trial in the training set was 3.55. The average minor criterion per clinical trial in the training data set was 7.3. Similarly, the testing dataset had an average count of 22.5 total eligibility criteria with a median of 20. We again saw there were more than double the amount that were classified as minor than were classified major. The average major criterion per clinical trial in the training set was 3.2. The average minor criterion per clinical trial in the training data set was 6.9. We annotated the study participant agreed upon major eligibility criteria from the initial interviews as well as from the follow-up surveys. We did not detect any new standardized major eligibility criteria in the follow-up survey. The instance count of the training and testing datasets are seen in Table 2 sorted by total dataset instances. Study participants completed a 5-point Likert scale to classify each major eligibility criterion as “Definitely Major”; “Possibly Major–Depending on the Context”; “Undecided”; “Possibly Minor”; and “Definitely Minor”. We report the mode from participant responses from the follow-up survey Likert scale in Table 2.

<table>
<thead>
<tr>
<th>Major Eligibility Criterion</th>
<th>Training Instances</th>
<th>Testing Instances</th>
<th>Total Dataset Instances</th>
<th>Likert Scale Mode for the Participant Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease staging</td>
<td>20</td>
<td>12</td>
<td>32</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>Prior treatment</td>
<td>17</td>
<td>8</td>
<td>25</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>Hormone receptor status</td>
<td>9</td>
<td>3</td>
<td>12</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>Gender</td>
<td>7</td>
<td>2</td>
<td>9</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>Histological confirmed diagnosis</td>
<td>6</td>
<td>2</td>
<td>8</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>HER2 status</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>Pre/Postmenopausal status</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>Comorbidities</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>Eligible to receive treatment</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>Diagnosis made by core needle biopsy</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>Age</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>Laboratory values</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>At least one breast available</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>Scheduled to receive treatment</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>Possibly Minor</td>
</tr>
<tr>
<td>Measurable disease</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Definitely Major</td>
</tr>
<tr>
<td>Future scheduled treatment</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Possibly Minor</td>
</tr>
<tr>
<td>Blood pressure</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>Language</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Undecided</td>
</tr>
<tr>
<td>Heart rate</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Possibly Major</td>
</tr>
<tr>
<td>Family history of breast cancer</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Possibly Minor</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Possibly Minor</td>
</tr>
<tr>
<td>Suitable to undergo MRI</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Possibly Minor</td>
</tr>
<tr>
<td>Have mobile phone</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>Definitely Minor</td>
</tr>
</tbody>
</table>

Figure 3 below displays major eligibility criteria as ranked by study participants in the final survey in decreasing order of significance\(^3\). The criteria are ranked by significance score, which was calculated by summing all participant scores for each criterion. The criteria received the following scores for each classification: “Definitely Major”=1; “Possibly Major–Depending on the Context”=2; “Undecided”=3; “Possibly Minor”=4; and “Definitely
The criterion with the lowest score is considered the most significant. For example, “Disease staging” had the lowest significance score because all participants identified it as “Definitely Major”.

![Figure 3. Major eligibility criteria measured with a Likert scale by participants in the follow-up surveys.](image)

**Discussion**

Our training dataset was considered a representative sample of interventional breast cancer clinical trials for three main reasons. First, our results showed that “Hormone receptor status” was identified as the third most common major eligibility criterion with 12 total instances in our dataset and second most significant major eligibility criterion after disease staging. This observation confirmed sample representativeness since the prevalence of hormone receptor positive breast cancer in the entire breast cancer population is approximately two-thirds of all reported cases\(^2\). Second, saturation was reached in our training dataset of major eligibility criteria instances since no new major eligibility criteria were detected in our testing dataset. Third, the sample represented over 41% of interventional breast cancer trials available at CUMC.

The results of our mixed-methods study prove the existence of major eligibility criteria. In addition, both the training and testing datasets had similar summary statistics. The average count of total eligibility criteria for both training and testing datasets was 19.7 and 22.5, respectively. The average count of major eligibility criteria for training and testing data sets was 3.55 and 3.2, respectively. Minor eligibility criteria are more prevalent in our sample of clinical trials than major criteria. This finding is significant because automating a system that needs to search for fewer machine-readable standardized criteria is more efficient and feasibly operational. In addition, this finding questions prior efforts to provide a complex knowledge representation of highly expressive concepts that gives equal weight to all eligibility criteria. Our study suggests that time could be saved by focusing on major criteria that were identified by study participants. It is possible that raters identify minor criteria more easily than major criteria, but an inescapable conclusion is that complex knowledge representations that give equal weight to all criteria are perhaps less useful than previously thought. This is an area that requires further exploration.

We have derived a standardized list of major eligibility criteria for breast cancer clinical trials at CUMC. We have confidence in our list being complete because we did not detect any new standardized major eligibility criteria in our testing dataset during the follow-up survey. While we focused exclusively on breast cancer, our study provides a mixed-methods approach for discovering major eligibility concepts in other diseases. This approach will be helpful in future projects since clinical trials lack standardized eligibility criteria. As a consequence, recognizing similar criteria across two or more clinical trials remains a challenge even in the same disease realm.

Our approach differs from previous research attempts to create eligibility criteria knowledge representations by avoiding highly expressive standardization and temporal restraints in favor of focusing on only the most user-relevant concepts. The granularity of the major eligibility criteria desiderata is coarser than other knowledge representations and more comparable to available electronic health record data. The coarseness of our standardized criteria is a very strong advantage for the usefulness of major eligibility criteria because it takes into account practical considerations, such as data availability in electronic health records and fitness to the clinic workflow. Our
focus group detailed how clinical research professionals use major eligibility criteria in their thought process while prescreening patients. Focus group data provide insight into building a clinical decision support system that is powered by major eligibility criteria to automatically suggest clinical trials that patients are potentially eligible for during clinical encounters.

We also sought to understand the rationale of major eligibility criteria. The study participants viewed HIV as less important for breast cancer clinical trial recruitment because it is rare among the target population. Other criteria, such as age, depend on context. Annotated transcripts from the initial interviews, the focus group session, and the follow-up survey provide evidence that prior treatment is significant, yet difficult to obtain. Thus, prior treatment is difficult to utilize in automatic prescreening. More advanced disease staging may increase the importance of some eligibility criteria such as laboratory results and prior treatment. As a consequence, clinical trials for early stage breast cancer could realize more of a benefit using major eligibility criteria for automated prescreening than studies for advanced disease stages.

The results of our mixed-methods study also identify the desiderata of major eligibility criteria for breast cancer clinical trials. While we utilized only breast cancer clinical trials, our desiderata could be scalable to other diseases. Furthermore, the desiderata could provide guidance about what features will create an optimal clinical trials recruitment tool for future clinical decision support systems. Since the desiderata terms are ranked by frequency, they provide weighted insight into how clinical researchers think while prescreening patients for clinical trials. For example, “Reliable” was the desiderata attribute that was most frequent to our major eligibility criteria followed by “Objective” and “Relevant to me”. The relatively highly ranked “Relevant to me” desiderata attribute suggests that a role-based analysis may provide additional insight into prescreening processes.

Limitations
Since the results are only from one academic medical center, we cannot speculate to the universality of the desiderata. We purposely included only breast cancer clinical trials in our dataset because previous research findings suggest that relevant eligibility criteria are disease specific. The applicability of our findings to other diseases is unknown. This study did not use any real patient data, instead relying on observing the opinions and thought processes of clinical research professionals using the “Think-aloud” protocol. As a consequence, we are unsure what effect a fragmented or incomplete patient electronic health record would have on influencing the major eligibility criteria desiderata. Finally, even though our study methods included qualitative and quantitative methods, our small sample size prohibits us from making definitive conclusions about breast cancer clinical trials in a larger context.

Future work
The scalability of the major eligibility criteria desiderata to other diseases and institutions is an unanswered research question that warrants additional exploration. Our desiderata may assist clinical researchers and informaticists to make informed decisions about prioritizing the standardization and collection of data in the electronic health record for clinical trial prescreening. 100% of study participants agreed disease stage was “Definitely Major” during the follow-up survey. This suggests that we should prioritize the capture of this information in a standardized, machine-readable way for efficient prescreening. Similarly, “Hormone receptor status” and “HER2 status” should be prioritized for standardized capture. Standardized major eligibility criteria could provide a basis for clinical decision support by triaging patients for clinical trial prescreening.

Future work should evaluate how major eligibility criteria are captured in electronic health records and how prescreening results can be best presented to clinical researchers in the clinic workflow. Our study has shown how major eligibility criteria might be used in an optimized workflow to screen patients for clinical trials. Automated prescreening could reduce human effort and expense by reducing the manual review of patient charts. If financial or other barriers prevent major eligibility criteria from being collected in a standardized way or even preclude an automated prescreening process, the benefits of major eligibility criteria still exist. For example, major eligibility criteria can still inform the development of prescreening methods. Additional evaluation of the desiderata and standardized major eligibility criteria will be helpful for understanding how our research can be utilized when prescreening patients in real life scenarios.

Conclusion
This study initially investigated how, when, and why clinical researchers use major eligibility criteria to triage patients for clinical trial prescreening. Our results demonstrate that major eligibility criteria exist for breast cancer clinical trials. Our identified desiderata can potentially inform future design of clinical decision support for clinical trial prescreening. Future research is warranted to test the generalizability of these results beyond breast cancer trials and consider real patient data in order to gain insight into how the major eligibility criteria desiderata is influenced by fragmented or incomplete patient records and what steps might be taken to mitigate those effects.
Acknowledgments

This study was sponsored by the U.S. National Library of Medicine grant R01LM009886 (PI: Weng) and U.S. National Center for Advancing Translational Science grant UL1 TR000040 (PI: Ginsberg).

References