Research and Applications

Research participation preferences as expressed through a patient portal: implications of demographic characteristics

Jihad S. Obeid,1 Azza Shoaibi,1 Jim C. Oates,2,3 Melissa L. Habrat,1 Chanita Hughes-Halbert,4 and Leslie A. Lenert1

1Biomedical Informatics Center, Medical University of South Carolina, Charleston, South Carolina, USA, 2Department of Medicine, Medical University of South Carolina, Charleston, SC, USA, 3Medical Service, Rheumatology Section, Ralph H. Johnson VA Medical Center, Charleston, SC, USA and 4Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, USA

Corresponding Author: Jihad S. Obeid, Medical University of South Carolina, Charleston, SC, USA (jobeid@musc.edu)

ABSTRACT

Objective: As patient portals are increasingly used for research recruitment, it is important to examine the demographic makeup of research registries that are populated via portals and the factors that influence participation in these registries.

Methods: We examined the response to a routine research preference questionnaire among patients who were enrolled in a patient portal at an academic health center and characterized the sub-population that responded and was tracked in a research preferences registry. We examined the factors that influence choices in two research preferences: future contact for research opportunities and biobanking of de-identified specimens.

Results: Out of 79,834 patients to whom the questionnaire was sent, 32% responded. Of those 74% agreed to future contact and 77% to the biobank preference. We found significantly lower odds of agreement in both preferences in minority populations, especially in the population >65 years of age when stratified by race. Individuals with higher comorbidity indexes had significantly higher odds for agreement.

Discussion: The disparities in volunteerism as expressed by agreement to future contact and willingness to participate in biobanking are exacerbated by lower levels of enrollment in the patient portal by minorities, especially in the oldest age group. Future work should examine other socioeconomic factors and the differences across age groups, sicker individuals, and payer categories.

Conclusion: Although patient portals can be more efficient for recruitment, researchers have to be cognizant of, and proactively address, potential biases when recruiting participants from these registries.

Key words: research preferences, patient portal, research recruitment

INTRODUCTION

Recruitment of participants into clinical trials has several challenges with financial and ethical ramifications.1–3 In the era of precision medicine research with lofty recruitment goals, there has been several novel approaches to reach out to potential participants through volunteer registries.4,5 As healthcare institutions move toward patient-centered care, personal health records (PHRs) play an increasing role with healthcare delivery.6–8 As a result, PHRs in the form of patient portals are being explored by many academic health centers as another venue for engaging patients in research for both patient reported outcomes and recruitment into clinical trials.9 In particular, patient portals that include a PHR tethered to the
Retention / Disposal and Use of Blood, Body Fluids, or Tissue.

I understand that any blood, body fluids or tissues normally removed from my body by MUSC Health, including its hospital(s), physicians and staff, in the course of any diagnostic procedures, surgery, or medical treatment that would otherwise be disposed of may be retained, and used for research, including research on the genetic material (DNA) or other information contained in those tissues or specimens. I acknowledge that such research by MUSC Health, may result in new inventions that may have commercial value and I understand that there are no plans to compensate me should this occur, regardless of the value of any such invention. I understand that any research using these leftover specimens or tissues will be done in a way that will not identify me. If I have questions, I should call (843) 792-8300 or visit http://www.muschealth.org/clinical-trials/

I AGREE to have my leftover blood, body fluids or tissue used for future research studies.

I do not agree to have my leftover blood, body fluids or tissue used for future research studies.

I am not ready to make a decision at this time.

Preference for Contact by Research Studies.

I understand I may be contacted about future research studies at MUSC for which I may be eligible. If I have questions, I should call (843) 792-8300 or http://www.muschealth.org/clinical-trials/

I AGREE to be contacted about future research studies.

I do not agree to be contacted about future research studies.

I am not ready to make a decision at this time.

Figure 1. The research preference questionnaire as expressed through a message via the patient portal includes the two preferences: (1) preference for allowing potential retention and use for de-identified body fluids or tissues specimens and (2) preference for contact about future research studies.

METHODS

The MUSC EHR-based research preferences registry includes responses from participants collected using a routine research preferences questionnaire via the patient portal. The details of the procedure used in constructing this registry have been previously described.14 The questionnaire includes preferences for two items: (1) preference for contact about future research studies (contact preference) and (2) preference for allowing potential retention and use for de-identified research of blood, body fluids, or tissues collected during routine clinical care that would otherwise be discarded (biobank preference). If patients choose to respond to the questionnaire, they are provided with three options to each of those two preferences: they may agree, disagree, or choose not to make a decision at this time (Figure 1). Also included in the questionnaire are links to detailed information about these two preferences and contact information for institutional representatives for further questions. The preferences registry serves as a virtual research volunteer registry that is tethered to the patient’s longitudinal health record.

We obtained approval from the MUSC Institutional Review Board (protocol # Pro00040823) for a waiver of consent and HIPAA authorization to conduct this study.

Data preparation

The study population included the patients enrolled in our patient portal who were invited to participate in the research preferences questionnaire, and the sub-population who responded to the questionnaire and were tracked in the research preferences registry during the period between December 2014 and May 2016. During this study period, all adults with an active account in the patient portal were invited to participate in the questionnaire. Therefore,
the study was limited to adults (18 years of age or older). Patients who were deceased at the time the questionnaire was sent were excluded. The variables collected for this study were limited to de-identified data from the EHR Research Data Warehouse (RDW). Demographic variables included gender, age in years at the time the questionnaire was sent, race, ethnicity, and marital status. Other variables included: Charlson Comorbidity Index (CCI)\(^ {18,19}\) pre-computed in our RDW and payer category. The outcome variables included: (1) status of response to the research preferences questionnaires (responded or not responded), (2) whether they agreed to the future contact, and (3) whether they agreed to the biobank preference. If patients had submitted two responses during the study period, only the latest response was considered. For the purpose of this analysis, the response to a research preference for both contact and biobank was categorized as agreed to the permission or decline agreement (a binary outcome). Disagreement and a non-decision were categorized as declining agreement. For comparative analysis, we examined the race and gender breakdown across our study populations and across a cross-section of the MUSC patient population 18 years of age or older at the time of this analysis in the RDW.

### Analysis

The data were analyzed using R software for statistical computing v3.4.3. The continuous variables age and CCI were converted to ordinal groups: age groups were categorized into 18–35, 36–50, 51–65, and >65; CCI groups were categorized into 0, 1, 2–3, 4–7, and >7. Pearson’s \( \chi^2 \) test was used to estimate the \( P \)-values across population characteristics. Multivariate logistic regression models were used to estimate the adjusted odds ratios (OR) for the binary outcomes of agreement in both contact and biobank preferences within those who responded to the questionnaire. We conducted a sensitivity analysis by using the propensity to respond to the questionnaire in the model for estimating the odds for agreement in both preferences. The propensity for response was predicted using a multivariate logistic regression model within the whole population to whom the questionnaire was sent. We also tested for interactions between race and other variables and performed a stratified analysis by race for the significant variables.

### RESULTS

The total number of individuals who were invited to participate in the questionnaire via the patient portal was 79,834. Of those, 25,768 (32%) responded to the invitation. The breakdown across the different variables in these two populations is shown in Table 1.

Of the 25,768 who responded, 18,892 (73.3%) agreed to future contact and 19,713 (76.5%) agreed to the biobank preference (Figure 2). Table 2 shows the population characteristics across both preferences.

There was significant attrition among minority patients (\( P < 0.001 \) across all phases) from enrollment in the patient portal (from 25.1% African American of the EHR adult patient population to 18.0% in the patient portal), response to the questionnaire (11.0%), and agreement in the research preferences (9.1% for contact and 8.1% for biobank) (Figure 3). The gender breakdown across the same populations shows a preponderance of females within the population of individuals enrolled in the patient portal compared to the EHR patient population ages 18 years and older (\( P < 0.001 \)). In contrast to the race breakdown, the gender trend reverses showing an increasing proportion of the males across the rest of the phases of the study (Figure 4).

The adjusted OR’s showed reduced levels of agreement in several population categories across both the contact and biobank

### Table 1. The population breakdown and percentages across the different factors with in each variable for the study populations

| Variable | Factor | Q. Sent (%) | Responded (%) |
|----------|--------|-------------|---------------|
| Gender   | Female | 52,129 (65.3) | 16,591 (64.4) |
|          | Male   | 27,705 (34.7) | 9,177 (35.6)  |
| Age      | 18–35  | 21,249 (26.6) | 4,437 (17.2)  |
|          | 36–50  | 19,091 (23.9) | 5,637 (21.9)  |
|          | 51–65  | 22,477 (28.2) | 8,849 (34.3)  |
|          | >65    | 17,017 (21.3) | 6,845 (26.6)  |
| Race     | White  | 61,448 (77)   | 21,882 (84.9) |
|          | African American | 14,336 (18)   | 2,842 (11)    |
|          | American Indian | 121 (0.2)      | 29 (0.1)      |
|          | Asian   | 853 (1.1)     | 202 (0.8)     |
|          | Other   | 1,874 (2.3)   | 441 (1.7)     |
|          | Unknown | 1202 (1.5)    | 372 (1.4)     |
| Ethnicity| Non-Hispanic | 74,762 (93.6) | 24,523 (95.2) |
|          | Hispanic | 889 (1.1)      | 225 (0.9)     |
|          | Unknown | 4183 (5.2)    | 1020 (4)      |
| Marital status | Married or significant other | 5086 (6.4) | 1743 (6.8) |
|          | Divorced or separated | 48,193 (60.4) | 17,532 (68) |
|          | Single  | 22,315 (28)   | 5301 (20.6)   |
|          | Widowed | 2497 (3.1)    | 787 (3.1)     |
|          | Unknown | 1743 (2.2)    | 405 (1.6)     |
| Payer category | Private or commercial | 46,176 (57.8) | 14,793 (57.4) |
|          | Medicaid | 5154 (6.5)    | 902 (3.5)     |
|          | Medicare | 21,973 (27.5) | 8556 (33.2)   |
|          | Military or Other Gov. | 2258 (2.8) | 706 (2.7) |
| CCI      | C_0     | 40,963 (51.3) | 11,564 (44.9) |
|          | C_1     | 15,004 (18.8) | 5104 (19.8)   |
|          | C_2–3   | 12,661 (15.9) | 4858 (18.9)   |
|          | C_4–7   | 6011 (7.5)    | 2214 (8.6)    |
|          | C_8+    | 5195 (6.5)    | 2028 (7.9)    |

\( \chi^2 \) test \( P \)-values <0.001 for all variables.

Abbreviations: Q. Sent: number of individuals who were invited to participate in the questionnaire via the portal; Responded: number of individuals who responded to the questionnaire; CCI: Charlson Comorbidity Index; C_0: CCI = 0; C_1: CCI = 1; C_2–3: CCI = 2 or 3; C_4–7: CCI between 4 and 7 inclusive; C_8+: CCI 8 or more.

![Figure 2. Breakdown among respondents to the future contact preference (A) and biobank preference (B). Declined — disagreed or chose not to make a decision.](image-url)
preferences (Table 3). The most significantly low OR was in the African American population for the contact preference (0.38) and the biobank preference (0.24). There were also significant reductions, albeit not as pronounced, in other race and ethnicity categories, for example Asian Americans, for both preferences and Hispanic individuals but only for the biobank preference. Interestingly, for marital status, single, divorced or separated, and widowed individuals, had low OR’s for agreement in both preferences compared to couples (married or significant other status), but only significantly for the contact preference. In contrast, the most important variables with positive impact on agreement in both preferences were higher categories of the CCI, i.e. sicker individuals tended to show higher volunteerism in research preferences (OR ranged 1.31–1.90 for contact and 1.26–2.38 for biobank). The only payer category that had a significant positive impact on both preferences was Medicare (OR 1.31–1.90) relative to the private or commercial payer category.

When we examined interactions of race with other variables, interactions with age and payer category stood out. A race-stratified analysis for age showed significantly reduced adjusted OR’s for age over 65 (with ages 18–35 as reference) in agreement to the biobank preference for both white (OR 0.80, 0.049) and African American (OR 0.48, 0.001) populations and in agreement to future contact only for African Americans (OR 0.49, 0.001). When stratified by race, payer category showed a significantly increased OR for agreement in both preferences in the self-pay payer category in African Americans. There was not such an effect in the White sub-population. See Table 4 for details.

**DISCUSSION**

Patient portals that are integrated with EHR data provide a unique opportunity for an informatics solution by matching interested patients with research protocol inclusion criteria based on their EHR phenotypes.\(^2\,\(^1\,\(^2\)\) However, researchers using this powerful tool should be aware of the advantages as well as the limitations. The results from our analysis bring to light several interesting observations.

Although the racial and gender breakdowns in the EHR adult patient population at MUSC are representative of those in Charleston County,\(^2\,\(^3\)\) the racial and gender breakdown in the patient portal populations deviates significantly from both the above. The disparities in race distributions in the patient portal are exacerbated, as less minority individuals respond to the research questionnaires, and even less volunteer to opt into research (Figure 3). In the context of

### Table 2. The population characteristics and percentages of agreement in both contact and biobank preferences within each factor for the individuals who responded to the questionnaire (total n = 25 768)

| Variable          | Factor                  | Contact preference | Biobank preference |
|-------------------|-------------------------|--------------------|--------------------|
|                   | Agreed                  | Declined           | Agreed             | Declined           |
| Total             | 18 892 (73.3%)          | 6876 (26.7%)       | 19 713 (76.5%)     | 6055 (23.5%)       |
| Gender            | Female                  | 11 863 (71.5%)     | 4728 (28.5%)       | 12 290 (74.1%)     | 4301 (25.9%)       |
|                   | Male                    | 7029 (76.6%)       | 2148 (23.4%)       | 7423 (80.9%)       | 1754 (19.1%)       |
| Age               | 18–35                   | 3023 (68.1%)       | 1414 (31.9%)       | 3235 (72.9%)       | 1202 (27.1%)       |
|                   | 36–50                   | 4055 (71.9%)       | 1582 (28.1%)       | 4181 (74.2%)       | 1456 (25.8%)       |
|                   | 51–65                   | 6580 (74.4%)       | 2269 (25.6%)       | 6791 (76.7%)       | 2058 (23.3%)       |
|                   | >65                     | 5234 (76.5%)       | 1611 (23.5%)       | 5506 (80.4%)       | 1339 (19.6%)       |
| Race              | White                   | 16 492 (75.4%)     | 5390 (24.6%)       | 17 402 (79.5%)     | 4480 (20.5%)       |
|                   | African American        | 1717 (60.4%)       | 1125 (39.6%)       | 1596 (56.2%)       | 1246 (43.8%)       |
|                   | American Indian         | 19 (65.5%)         | 10 (34.5%)         | 17 (58.6%)         | 12 (41.4%)         |
|                   | Asian                   | 116 (57.4%)        | 86 (42.6%)         | 130 (64.4%)        | 72 (35.6%)         |
|                   | Other                   | 303 (68.7%)        | 138 (31.3%)        | 315 (71.4%)        | 126 (28.6%)        |
|                   | Unknown                 | 245 (63.9%)        | 127 (34.1%)        | 253 (68%)          | 119 (32%)          |
| Ethnicity         | Non-Hispanic            | 18 031 (73.5%)     | 6492 (26.5%)       | 18 827 (76.8%)     | 5696 (23.2%)       |
|                   | Hispanic                | 166 (73.8%)        | 59 (26.2%)         | 157 (69.8%)        | 68 (30.2%)         |
|                   | Unknown                 | 695 (68.1%)        | 325 (31.9%)        | 729 (71.5%)        | 291 (28.5%)        |
| Marital status    | Married or significant other | 1337 (77.9%)  | 386 (22.1%)       | 1323 (75.9%)       | 420 (24.1%)        |
|                   | Divorced or separated   | 12 996 (74.1%)     | 4536 (25.9%)       | 13 679 (78%)       | 3853 (22%)         |
|                   | Single                  | 3722 (70.2%)       | 1579 (29.8%)       | 3824 (72.1%)       | 1477 (27.9%)       |
|                   | Widowed                 | 555 (70.5%)        | 232 (29.5%)        | 608 (77.3%)        | 179 (22.7%)        |
|                   | Unknown                 | 262 (64.7%)        | 143 (35.3%)        | 279 (69.8%)        | 126 (31.2%)        |
| Payer category    | Private or commercial   | 10 543 (71.3%)     | 4250 (28.7%)       | 11 049 (74.7%)     | 3744 (25.3%)       |
|                   | Medicaid                | 652 (72.3%)        | 250 (27.7%)        | 630 (69.8%)        | 272 (30.2%)        |
|                   | Medicare                | 6622 (77.4%)       | 1934 (22.6%)       | 6903 (80.7%)       | 1653 (19.3%)       |
|                   | Military or Other Gov.  | 498 (70.5%)        | 208 (29.5%)        | 544 (77.1%)        | 162 (22.9%)        |
|                   | Self-pay                | 577 (71.1%)        | 234 (28.9%)        | 587 (72.4%)        | 224 (27.6%)        |
| CCI               | C_0                     | 8029 (69.4%)       | 3335 (30.6%)       | 8465 (73.2%)       | 3099 (26.8%)       |
|                   | C_1                     | 3774 (73.9%)       | 1330 (26.1%)       | 3865 (75.7%)       | 1239 (24.3%)       |
|                   | C_2–3                   | 3733 (76.8%)       | 1125 (23.2%)       | 3857 (79.4%)       | 1001 (20.6%)       |
|                   | C_4–7                   | 1703 (76.9%)       | 511 (23.1%)        | 1773 (80.1%)       | 441 (19.9%)        |
|                   | C_8+                    | 1633 (81.5%)       | 375 (18.5%)        | 1753 (86.4%)       | 273 (13.6%)        |

\(^2\) test P-values <0.001 for all variables for both preferences.

Abbreviations: CCI: Charlson Comorbidity Index; C_0: CCI = 0; C_1: CCI = 1; C_2–3: CCI = 2 or 3; C_4–7: CCI between 4 and 7 inclusive; C_8+: CCI 8 or more.
this discussion, we refer to volunteerism as volunteering to be contacted for future research or donating de-identified specimens for biobanking. Our multivariate analysis shows race as being the most important variable in determining low research volunteerism. The results show significantly lower odds of opting in for both future research contact and biobanking in minority populations. Non-response bias aggravates the problem even further. Adding the probability of response to the questionnaire into logistic regression model for opting-in reduces the odds for participation in research for African Americans even more. Because our analysis was limited to de-identified data that were available in the RDW, we were not able to examine the associations between research preference and distrust or other factors that have been associated with low minority participation in research. However, there is significant evidence in the literature to support reasons for reduced enrollment of African Americans in biobanks including lack of trust due to fear of research exploitation, rooted in past unethical medical research involving African Americans,24–27 as well as failure of adequate recruitment
by the research community.28 The associations between distrust and attitudes about participating in medical research and volunteering to participate in a hospital-based research registry should be examined in future studies.

In contrast, males in both white and African American populations had higher odds of agreement in both preferences relative to females, even though our patient portal population seems to have a significantly higher proportion of females (65.3%) (Figure 4). Another interesting finding is the increased volunteerism evidenced by the significantly higher odds of agreement in both preferences (but more so for biobanking) in patients who are ill, using the Charlson index as proxy (Table 3). The impact of age varies by race, as shown in the stratified analysis (Table 4). Particularly notable is the lower volunteerism in either preference in African Americans older than 65 years of age (P < 0.001) in the stratified analysis when compared to younger individuals. This could represent the impact of a generational memory and the distrust issue noted above. This could also be an encouraging finding and possibly an indication that future generations may have a smaller racial gap in volunteerism for research. However, further analysis is needed before we can make such an assertion.

In order to overcome some of the patient portal disparities, other modes of registry enrollment should be sought. For example, at MUSC we have begun providing the same registry questionnaire during in-clinic patient visits. Analysis of the impact of this new method of presenting the registry questionnaire on the opt-in rate is underway. Other approaches should include targeted educational and patient engagement campaigns,29 healthcare providers engagement in research and proactively address, potential selection bias when using such recruitment registries especially those involving opt-in approaches through EHR patient portals. Several methodological and statistical

### Table 3. The adjusted odds ratios (OR) and 95% confidence intervals (CI) for the binary outcomes of agreement in both the contact and bio-bank preferences

| Variable          | Factor                  | Contact OR (CI) | P-value | Biobank OR (CI) | P-value |
|-------------------|-------------------------|-----------------|---------|-----------------|---------|
| Gender            | Female                  | Ref             |         | Ref             |         |
|                   | Male                    | 1.11 (0.99–1.23) | 0.072   | 1.19 (1.06–1.33) | 0.003   |
| Age               | 18–35                   | Ref             |         | Ref             |         |
|                   | 36–50                   | 1.27 (1.02–1.57) | 0.032   | 1.13 (0.91–1.4)  | 0.268   |
|                   | 51–65                   | 1.41 (0.93–2.15) | 0.106   | 1.26 (0.83–1.93) | 0.276   |
|                   | >65                     | 1.17 (0.76–1.78) | 0.478   | 1.10 (0.71–1.68) | 0.677   |
| Race              | White                   | Ref             |         | Ref             |         |
|                   | African American        | 0.38 (0.26–0.56) | <0.001 | 0.24 (0.16–0.35) | <0.001 |
|                   | American Indian         | 0.45 (0.2–1.05)  | 0.064   | 0.24 (0.11–0.55) | <0.001 |
|                   | Asian                   | 0.40 (0.27–0.6)  | <0.001 | 0.39 (0.26–0.58) | <0.001 |
|                   | Other                   | 0.63 (0.46–0.86) | 0.004   | 0.59 (0.43–0.81) | 0.001   |
|                   | Unknown                 | 1.05 (0.69–1.59) | 0.833   | 0.98 (0.64–1.5)  | 0.926   |
| Ethnicity         | Non-Hispanic            | Ref             |         | Ref             |         |
|                   | Hispanic                | 1.22 (0.88–1.7)  | 0.224   | 0.81 (0.59–1.11) | 0.190   |
|                   | Unknown                 | 0.79 (0.62–1.02) | 0.070   | 0.74 (0.57–0.96) | 0.021   |
| Marital status    | Married or significant other | Ref         |         | Ref             |         |
|                   | Divorced or separated   | 0.79 (0.7–0.91)  | <0.001 | 1.02 (0.89–1.16) | 0.770   |
|                   | Single                  | 0.76 (0.66–0.87) | <0.001 | 0.89 (0.77–1.02) | 0.086   |
|                   | Widowed                 | 0.52 (0.39–0.7)  | <0.001 | 0.76 (0.56–1.03) | 0.082   |
|                   | Unknown                 | 0.54 (0.4–0.73)  | <0.001 | 0.67 (0.49–0.91) | 0.010   |
| Payer category    | Private or commercial   | Ref             |         | Ref             |         |
|                   | Medicaid                | 1.20 (0.95–1.52) | 0.127   | 0.97 (0.77–1.23) | 0.802   |
|                   | Medicare                | 1.31 (1.18–1.44) | <0.001 | 1.27 (1.15–1.41) | <0.001 |
|                   | Military or Other Gov.  | 0.93 (0.78–1.11) | 0.412   | 1.08 (0.89–1.3)  | 0.437   |
|                   | Self-pay                | 0.90 (0.65–1.26) | 0.552   | 0.76 (0.54–1.07) | 0.115   |
| CCI               | C_0                     | Ref             |         | Ref             |         |
|                   | C_1                     | 1.31 (1.15–1.49) | <0.001 | 1.26 (1.1–1.44)  | <0.001 |
|                   | C_2–3                   | 1.53 (1.28–1.83) | <0.001 | 1.60 (1.33–1.92) | <0.001 |
|                   | C_4–7                   | 1.46 (1.24–1.72) | <0.001 | 1.56 (1.31–1.85) | <0.001 |
|                   | C_8+                    | 1.90 (1.38–2.28) | <0.001 | 2.38 (1.96–2.89) | <0.001 |

Significant P-values are bolded. Ref indicates factors used as reference.

Abbreviations: CCI: Charlson Comorbidity Index; C_0: CCI = 0; C_1: CCI = 1; C_2–3: CCI = 2 or 3; C_4–7: CCI between 4 and 7 inclusive; C_8+: CCI 8 or more.
approaches have been described to reduce volunteer or self-selection bias during analysis including, but not limited to, generalization of inverse probability weighting or stratified analyses.32

Limitations and future directions
This analysis is not without limitations. Other than payer category, socioeconomic factors have not been considered due to the lack of such data in our RDW. Future work involving surveys may help tease out the differences in volunteerism and the reasons for variations across different age groups, CCI groups, and payer categories. Although our registry population is fairly large, this study represents data from a single institution in a specific region of the country. Collaborative studies with other institutions using similar registry practices need to be conducted in order to validate the transferability of our findings.

Although this work examines willingness to participate in the research registry, we cannot confirm that this translates into actual participation in clinical trials. We assume that individuals who are open to be contacted for research will be more likely to volunteer during future recruitment efforts into research protocols for which they are found to be eligible. A future analysis using data from our registry will be necessary to confirm this hypothesis.

CONCLUSION
An EHR-integrated volunteer registry via the patient portal lends itself to a unique informatics solution for matching willing potential participants with research protocols. However, this solution has its limitations. Our results show significant differences in patient portal membership as well as opt-in rates based on several criteria including race, gender, age, and health status. There needs to be further investigation of the causes of disparities across all these factors, the most pronounced of which is race. Patient and community engagement campaigns in under-represented populations may help allay some of these disparities. In the meantime, researchers have to be aware of these limitations when recruiting from such registries.

FUNDING
This work was supported in part by the NIH National Center for Advancing Translational Sciences (NCATS) through Grant Number UL1 TR001450 and the National Institute on Minority Health and Health Disparities (NIMHD) grant U54MD010706.

Conflict of interest statement. None declared.

ACKNOWLEDGMENTS
We would like to thank Dr Paul Nietert at MUSC for his advice on statistical analysis.

CONTRIBUTORSHIP STATEMENT
JSO, JCO, and LAL conceived and designed the study. MLH and AS collected the data. JSO and AS analyzed the data, designed the statistical method, and drafted the article. CHH and LAL critically revised the article for important intellectual content. All coauthors contributed to writing, proofreading, and editing the manuscript.

REFERENCES
1. Sung NS, Crowley WF, Genel M, et al. Central challenges facing the national clinical research enterprise. JAMA 2003; 289 (10): 1278–87.
2. Gul RB, Ali PA. Clinical trials: the challenge of recruitment and retention of participants. J Clin Nurs 2010; 19 (1-2): 227–33.

3. Cooley ME, Sarna L, Brown JK, et al. Challenges of recruitment and retention in multisite clinical research. Cancer Nurs 2003; 26 (5): 376–84; quiz 385–6.

4. Harris PA, Scott KW, Lebo L, Hassan N, Lightner C, Pulley J. ResearchMatch: a national registry to recruit volunteers for clinical research. Acad Med J Med 2012; 87 (1): 66–73.

5. Sankar PL, Parker LS. The Precision Medicine Initiative's All of Us Recruitment Program: an agenda for research on its ethical, legal, and social issues. Genet Med 2017; 19 (7): 743–50.

6. Kupchunas WR. Personal health record: new opportunity for patient education. Orthop Nurs 2007; 26 (3): 185–91; quiz 192–3.

7. Detmer D, Bloomrosen M, Raymond B, Tang P. Integrated personal health records: transformative tools for consumer-centric care. BMJ Med Inform Decis Mak 2008; 8 (1): 45.

8. Irizarry T, DeVito Dabbs A, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015; 17 (6): e148.

9. Obeid JS, Beskow LM, Rape M, et al. A survey of practices for the use of electronic health records to support research recruitment. J Clin Transl Sci 2017; 1 (4): 246–52.

10. Aronson SJ, Rehm HL. Building the foundation for genomics in precision medicine. Nature 2015; 526 (7573): 336–42.

11. Silvestre A-L, Sue VM, Allen JY. If you build it, will they come? The Kaiser Permanente model of online health care. Health Aff (Millwood) 2011; 30 (1): 173–82.

12. Obed JS, Gerken K, Madathil KC, et al. Development of an electronic research permissions management system to enhance informed consents and capture research authorizations data. AMIA Summits Transl Sci Proc 2013; 2013: 189–93.

13. Sanderson IC, Obeid JS, Madathil KC, et al. Managing clinical research permissions electronically: a novel approach to enhancing recruitment and managing consents. Clinical Trials 2013; 10 (4): 604–11.

14. Marshall EA, Oates JC, Shooabi A, et al. A population-based approach for implementing change from opt-out to opt-in research permissions. PLoS One 2017; 12 (4): e0168223.

15. Peacock S, Reddy A, Leveille SG, et al. Patient portals and personal health information online: perception, access, and use by US adults. J Am Med Inform Assoc JAMIA 2017; 24 (e1): e173–7.

16. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. Disparities in enrollment and use of an electronic patient portal. J Gen Intern Med 2011; 26 (10): 1112–6.

17. Sarkar U, Karter AJ, Liu JY, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc JAMIA 2011; 18 (3): 318–21.

18. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987; 40 (5): 373–83.

19. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005; 43 (11): 1130–9.

20. R Core Team. R: A Language and Environment for Statistical Computing. [Internet]. 2017. https://www.r-project.org/. Accessed February 9, 2018.

21. Chute CG, Pathak J, Savova GK, et al. The SHARPn project on secondary use of Electronic Medical Record data: progress, plans, and possibilities. AMIA Annu Symp Proc AMIA JAMIA 2011; 2011: 248–56.

22. Shivade C, Raghavan P, Foster-Lussier E, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. J Am Med Inform Assoc JAMIA 2014; 21 (2): 221–30.

23. U.S. Census Bureau QuickFacts: UNITED STATES [Internet]. https://www.census.gov/quickfacts/fact/table/US/PST045217 Accessed March 25, 2018.

24. Corbie-Smith G, Thomas SB, St George DMM. Distrust, race, and research. Arch Intern Med 2002; 162 (21): 2458–63.

25. McDonald JA, Vadaparampil S, Bowen D, et al. Intentions to donate to a biobank in a national sample of African Americans. Public Health Genomics 2014; 17 (3): 173–82.

26. Byrd GS, Edwards CL, Kelkar VA, et al. Recruiting intergenerational African American males for biomedical research studies: a major research challenge. J Natl Med Assoc 2011; 103 (6): 480–7.

27. Shavers-Hornaday VL, Lynch CF, Burmeister LF, Torner JC. Why are African Americans under-represented in medical research studies? Impediments to participation. Ethn Health 1997; 2 (1-2): 31–45.

28. Hagiwara N, Berry-Bobovski I, Francis C, Ramsey L, Chapman RA, Albrecht TL. Unexpected findings in the exploration of African American underrepresentation in biospecimen collection and biobanks. J Canc Educ 2014; 29 (3): 580–7.

29. Coulter A, Ellins J. Effectiveness of strategies for informing, educating, and involving patients. BMJ 2007; 335 (7609): 24–7.

30. Unertl KM, Schaeftbauer CL, Campbell TR, et al. Integrating community-based participatory research and informatics approaches to improve the engagement and health of underserved populations. J Am Med Inform Assoc JAMIA 2016; 23 (1): 60–73.

31. Oramasionwu CU, Vitko ML, Johnson TL, Sage AJ, Chung AE, Bailey SC. Assessing the quality and suitability of multimedia resources to assist patients with patient portal use. Am J Med Qual off Qual 2017; 32 (4): 453–5.

32. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004; 15 (5): 615–25.