How Are the Arts and Humanities Used in Medical Education? Results of a Scoping Review

Tracy Moniz, MA, PhD, Maryam Golafshani, MA, Carolyn M. Gaspar, MSc, Nancy E. Adams, MLS, EdD, Paul Haidet, MD, MPH, Javeed Sukhera, MD, PhD, Rebecca L. Volpe, PhD, Claire de Boer, MS, and Lorelei Lingard, MA, PhD

Abstract

Purpose
Although focused reviews have characterized subsets of the literature on the arts and humanities in medical education, a large-scale overview of the field is needed to inform efforts to strengthen these approaches in medicine.

Method
The authors conducted a scoping review in 2019 to identify how the arts and humanities are used to educate physicians and interprofessional learners across the medical education continuum in Canada and the United States. A search strategy involving 7 databases identified 21,985 citations. Five reviewers independently screened the titles and abstracts. Full-text screening followed (n = 4,649). Of these, 769 records met the inclusion criteria. The authors performed descriptive and statistical analyses and conducted semistructured interviews with 15 stakeholders.

Results
The literature is dominated by conceptual works (n = 294) that critically engaged with arts and humanities approaches or generally called for their use in medical education, followed by program descriptions (n = 255). The literary arts (n = 197) were most common. Less than a third of records explicitly engaged theory as a strong component (n = 230). Of descriptive and empirical records (n = 424), more than half concerned undergraduate medical education (n = 245). There were gaps in the literature on interprofessional education, program evaluation, and learner assessment. Programming was most often taught by medical faculty who published their initiatives (n = 236). Absent were voices of contributing artists, docents, and other arts and humanities practitioners from outside medicine. Stakeholders confirmed that these findings resonated with their experiences.

Conclusions
This literature is characterized by brief, episodic installments, privileging a biomedical orientation and largely lacking a theoretical frame to weave the installments into a larger story that accumulates over time and across subfields. These findings should inform efforts to promote, integrate, and study uses of the arts and humanities in medical education.

Over the past 30 years, medical education has increasingly embraced the arts and humanities as a teaching modality, from using visual art to teach observation skills to using literature to promote perspective taking. A diverse scholarly community has formed around this effort, and a national initiative is underway in the United States to support consistent, effective uptake of arts- and humanities-based approaches along the continuum of medical education. A growing base of smaller studies describes (and debates) the nature and impact of arts and humanities curricula in medicine. Many suggest that learning experiences that integrate the arts and humanities may lead to a variety of important learning outcomes, including skills-based outcomes like teamwork, and that some medical education challenges, such as teaching empathy, may be best addressed through engagement with the arts and humanities. Others argue that evidence is lacking to support the use of these disciplines in medical training—that educators have failed to articulate what the arts and humanities can achieve for medical education.

Amid such debates, the scholarly community lacks a large-scale, systematic overview of the literature that might advance the field and inform national efforts to strengthen arts and humanities curricula in medicine. The Association of American Medical Colleges (AAMC) is leading one such effort to put forth an arts and humanities foundation for the education of future physicians and to promote the integration of these disciplines throughout the medical education continuum. To inform this effort, this review seeks to describe how educators are using the arts and humanities and what opportunities and obstacles remain to support the integration of these disciplines along the medical education continuum.

Method
We conducted a scoping review commissioned by the AAMC to address the following research question: How and why are the arts and humanities being used to educate physicians and interprofessional learners across the continuum of medical education? We followed Arksey and O’Malley’s 5-stage scoping review methodology to review, summarize, and synthesize the literature on this topic. We also included the sixth scoping review stage of stakeholder consultations. Given the 12-month schedule of the AAMC-commissioned work, our scoping review design was also informed by rapid review methodology. Such methods simplify or omit components of the review process.
to produce timely information, such as limiting the search by publication date and language and excluding gray literature.\textsuperscript{11,12}

**Inclusion and exclusion criteria**

We considered a range of art forms and humanities subjects used in pursuit of medical education goals: literature, creative writing, reflective writing, narrative medicine, film and television, theater and drama, visual art, visual thinking strategies, graphic novels and comics, music, dance, theology, philosophy (excluding medical ethics unless explicitly taught using the arts and/or humanities), history (excluding the history of medicine unless a history of the arts and/or humanities in medical education was included), classics, women and gender studies, and critical theory and cultural studies. We included English-only results and restricted results by date of publication. We ultimately included records published since 1991—when K.M. Hunter’s foundational book *Doctors' Stories: The Narrative Structure of Medical Knowledge*\textsuperscript{13} was published—and added foundational historical works identified through discussion among the research team.

We included qualitative and quantitative research as well as descriptive and conceptual papers, research about elective and required experiences from premedical education through continuing medical education, and research about programs for physicians or physician learners, including those programs with interprofessional learners. We excluded gray literature\textsuperscript{11} as well as empirical and descriptive records about programming outside the United States or Canada. This decision reflected criteria outlined by the AAMC in commissioning the work, as its primary interest and its membership comprise academic medicine institutions in the United States and Canada. We did include conceptual pieces from other countries that were foundational works in the field and/or that spoke broadly to the arts and humanities in medical education.

**Data collection**

Five authors (T.M., M.G., C.M.G., N.E.A., and L.L.) developed a search strategy to identify records on the uses of the arts and humanities in medical education. One author (N.E.A.), a research librarian, implemented the search in May and June 2019 across 7 databases: PubMed, ERIC, CINAHL, Cochrane Library, Web of Science, PsycInfo, and EMBASE. The search string for PubMed is included in Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/B94. The search string was translated into the syntax and vocabulary of each additional database. Four authors (T.M., M.G., C.M.G., and J.S.) and 1 contributor (see Acknowledgments) screened the titles and abstracts of the resulting records (n = 21,985), applying the inclusion and exclusion criteria. Five authors (T.M., M.G., C.M.G., R.L.V., and C.d.B.) and 1 contributor (see Acknowledgments) completed the full-text screening (n = 4,649). At both screening stages, 2 trained reviewers screened each record using a screening guide we developed to reflect the inclusion and exclusion criteria. At each level of screening, 3 authors (T.M., M.G., and C.M.G.) piloted the guide, met regularly to discuss the process, and iteratively revised the guide as needed. Discrepancies were resolved by a third reviewer. In the end, 769 records met the inclusion criteria. A flowchart of the study selection process is provided in Figure 1.

We did not hand-search the bibliographies of the included records, a decision supported by rapid review procedures.\textsuperscript{12} However, as a test, we hand-searched the bibliographies of a random sample of 5% of the included records to see how many extra studies this process would yield for inclusion. We set a threshold of 10% for records missed (i.e., if we had missed more than 10% of the records we identified in this secondary search, we would hand-search the bibliographies of all included records). Of the sample, we identified 76 relevant records for inclusion. Of these, only 6 were not already included in our data set, which was within the acceptable threshold.

**Data charting and analysis**

This scoping review involved descriptive and statistical analyses of the selected records. For the descriptive analysis, we used methods from content and thematic analysis\textsuperscript{14} to extract and chart the features of the included records. We collectively decided to extract demographic information from all records, including publication type, research type, interprofessional or not, explicitly framed by theory or not, intentional use of the term “arts” or “humanities” or not, and art form and/or humanities subject. Where relevant, we also extracted information about curricular features, including medical learner level, elective or required, educational setting, instructor profile, program evaluation or not, and learner assessment or not. Following an iterative process,\textsuperscript{9,10,15} 4 authors (T.M., M.G., C.M.G., and L.L.) developed, piloted, and updated a data charting form to determine if the approach to data extraction was consistent with the research question and purpose.\textsuperscript{10} These same 4 authors independently charted the data using the descriptive-analytic method,\textsuperscript{9,10} meeting regularly to discuss the process.

We conducted a statistical analysis to determine whether there were significant relationships between select variables that were analyzed in the included records. We used chi-square tests to explore associations between the following variables: type of publication, type of research, and integration of theory or theoretical framework. We explored these variables in response to patterns emerging in the literature over the course of our data extraction and analysis. For example, we observed that journal articles (the dominant publication type in the included records) seemed less likely to integrate theory than books, but we wanted to determine whether the pattern we observed was statistically significant before making inferences about the implications for cohesion or shared foundations in the field. Following each chi-square test, we assessed adjusted residuals to further explore relationships between specific categories of variables; absolute residuals of 3 or greater were considered relevant.\textsuperscript{16,17} The *P* value threshold for significance was set at *P* < .01 to reduce the risk of type 1 error.

**Stakeholder interviews**

Alongside these analyses and informed by them, we conducted the sixth stage of the scoping review methodology: consultation with stakeholders. This stage served both as a key knowledge translation component and as a way to elaborate on the patterns and gaps identified in our review synthesis.\textsuperscript{10} We collaboratively developed a semistructured interview guide and a list of prospective interviewees based on both their reputation in the field and the descriptive findings of the scoping
Review, which pointed to missing voices in the literature (e.g., medical learners and artists). One author (P.H.) conducted the interviews, which were recorded, transcribed, and analyzed for themes. Participants were a purposive sample drawn from authors of seminal articles and books in the field, educators leading arts- and humanities-based curricula, leaders and administrators supporting curricula, artists or docents, and medical learners. Interviews were conducted in 2 rounds—the first round with two-thirds of the stakeholders was held after the initial data synthesis and the second round with the remaining stakeholders was held after further analysis and discussion of the findings.

Reflection on the scoping review team
This scoping review was conceptualized by a large team (see Acknowledgments). A core team of the coauthors on this article contributed to the initial conceptualization and participated substantially in at least one of the following stages: search, synthesis, or stakeholder interviews. This core team included 5 Canadian and 4 American members who brought a variety of relevant perspectives to the work. Two were medical or health learners (M.G. and C.M.G.), 2 were academic clinicians (P.H. and J.S.), 3 were nonclinician academic researchers (T.M., R.L.V., and L.L.), and 1 was a librarian (N.E.A.). Three (P.H., R.L.V., and C.d.B.) were medical educators who use the arts and humanities in their teaching, and 2 (P.H. and L.L.) had been authors on previously published scoping reviews of the field. Our perspectives and experiences influenced the work in important ways, particularly our analytical approaches, which reflect our familiarity with qualitative and discourse analysis techniques and our intimate knowledge of—and decision to build upon—the results of recent reviews.

The University Research Ethics Board at Mount Saint Vincent University and the Institutional Review Board at Penn State College of Medicine reviewed and cleared or exempted this work.

Results
Here, we present a synthesis of the published scholarship in the field, based on our descriptive and statistical analyses and on the stakeholder interviews. Given the large data set (N = 769), we cite selected, rather than all, records in a given category to illustrate our findings. A complete list of the included records is available in Supplemental Digital Appendix 2 at http://links.lww.com/ACADMED/B94.
Descriptive analysis

Most records were journal articles (610, 79%), followed by book chapters (144, 19%), books (10, 1%), and dissertations or theses (5, 1%). Conceptual (294, 38%) and descriptive (255, 33%) pieces were most common, followed by empirical studies (169, 22%) and reviews (51, 7%). Of the conceptual records, 171 (58%) were substantive theoretical contributions, 19 pieces that demonstrated critical, theoretical, or philosophical engagement with ideas or methods in the field; the remainder (123, 42%) served as a general call for the use of the arts and humanities in medical education. Of the empirical records, nearly half were qualitative studies (77, 45%), a third were studies that reported both qualitative and quantitative results (60, 36%), and a fifth were quantitative studies (32, 6%). Of the reviews, a third were described by their authors as literature reviews (17, 33%), with fewer described as systematic reviews (4, 8%), scoping reviews (3, 6%), and narrative reviews (3, 6%). Nearly half of the reviews were categorized as “other” (24, 47%), which included those articles described as a “review” where the authors did not further specify the type of review or where they provided a critical appraisal of a particular work, such as a book or film, for a medical education objective.

Only 73 records (10%) described programming that involved both medical learners and learners from other health professions. Ten records (1%) lacked sufficient information for us to determine whether the context was interprofessional, and the remaining 686 records (89%) focused on the education of physicians alone.

Records most often described the arts and humanities or medical/health humanities generally (170, 22%), either by not specifying a particular form or by invoking multiple forms (e.g., music, film, literature, dance). Our attempt to categorize all records according to whether the authors were intentional in their use of the terms “arts” or “humanities” was ultimately unsuccessful. Only rarely were authors explicit in defining these terms; more often they used them matter-of-factly without definition or exploration. Because of this, when coding the data, we relied on subjective interpretation of each record to assess the authors’ intentionality, resulting in our inability to achieve consistency among coders and our decision to abandon this categorization as unreliable.

Table 1 presents the distribution of the art forms and humanities subjects in the records we reviewed. Categories were not mutually exclusive, as a given record could focus on more than 1 art form and/or humanities subject. Categories were also not always self-evident, as authors used varied terminology within a category (e.g., reflection, reflective writing, reflective practice). In such cases, we used our judgment based on reading the full record to determine what the authors intended by a particular term.

Of all included records, 230 (30%) were explicitly framed by theory. Of these, more than half were conceptual pieces (128, 56%). Forty-four (19%) empirical and

Table 1

<table>
<thead>
<tr>
<th>Category and modality</th>
<th>Total records, no. (% of 769)</th>
<th>Records on modality, no. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Literature</strong></td>
<td>197 (26)</td>
<td>N/A</td>
</tr>
<tr>
<td>Reflective writing</td>
<td>119 (15)</td>
<td>N/A</td>
</tr>
<tr>
<td>Narrative medicine</td>
<td>86 (11)</td>
<td>N/A</td>
</tr>
<tr>
<td>Other writing*</td>
<td>61 (8)</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Visual art</strong></td>
<td>82 (11)</td>
<td></td>
</tr>
<tr>
<td>Observe</td>
<td>93 (48)</td>
<td></td>
</tr>
<tr>
<td>Create</td>
<td>22 (27)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>16 (19)</td>
<td></td>
</tr>
<tr>
<td>Unspecified</td>
<td>5 (6)</td>
<td></td>
</tr>
<tr>
<td><strong>Theater and drama</strong></td>
<td>70 (9)</td>
<td></td>
</tr>
<tr>
<td>Observe or read</td>
<td>22 (31)</td>
<td></td>
</tr>
<tr>
<td>Perform</td>
<td>24 (34)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>20 (29)</td>
<td></td>
</tr>
<tr>
<td>Unspecified</td>
<td>4 (6)</td>
<td></td>
</tr>
<tr>
<td><strong>Film and television</strong></td>
<td>67 (9)</td>
<td></td>
</tr>
<tr>
<td>Observe</td>
<td>62 (84)</td>
<td></td>
</tr>
<tr>
<td>Create</td>
<td>1 (2)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>2 (3)</td>
<td></td>
</tr>
<tr>
<td>Unspecified</td>
<td>1 (2)</td>
<td></td>
</tr>
<tr>
<td><strong>Other art forms</strong></td>
<td>32 (4)</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Music</strong></td>
<td>21 (3)</td>
<td></td>
</tr>
<tr>
<td>Listen</td>
<td>10 (48)</td>
<td></td>
</tr>
<tr>
<td>Create</td>
<td>5 (24)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>3 (14)</td>
<td></td>
</tr>
<tr>
<td>Unspecified</td>
<td>3 (14)</td>
<td></td>
</tr>
<tr>
<td><strong>Comics and graphic novels</strong></td>
<td>17 (2)</td>
<td></td>
</tr>
<tr>
<td>Read</td>
<td>10 (59)</td>
<td></td>
</tr>
<tr>
<td>Create</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>5 (29)</td>
<td></td>
</tr>
<tr>
<td>Unspecified</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td><strong>Philosophy</strong></td>
<td>17 (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>History</td>
<td>6 (1)</td>
<td>N/A</td>
</tr>
<tr>
<td>Religion</td>
<td>5 (1)</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Visual thinking strategies</strong></td>
<td>11 (1)</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: N/A, not applicable.

*Categories are not mutually exclusive.

*Examples of other writing include poetry and creative writing.

*Examples of other art forms include photography, podcasts, and TED Talks.
44 (19%) descriptive records explicitly grounded the work in theory (14, 6%).

We further analyzed the descriptive (255) and empirical (169) records (combined 424) for features of curricular programming. More than half of these records concerned undergraduate medical education (245, 58%), followed by postgraduate medical education (73, 17%), premedical education (20, 5%), and continuing medical education (13, 3%). At times, records focused on learners at multiple levels (63, 15%), with the most common of these being activities that spanned undergraduate to continuing medical education (34, 54%) or undergraduate to postgraduate medical education (17, 27%). Some records did not contain sufficient data to determine learner level (10, 2%).

About as many records described required programming (104, 24%) as elective programming (103, 24%), and 75 records (18%) focused on extracurricular programming. Twenty-eight records (7%)—all program descriptions—described a mix of required, elective, and extracurricular programming across the educational continuum. Nearly a third of records lacked sufficient information to determine the nature of program participation (114, 27%).

Programming most often took place in the classroom (178, 42%). To a lesser extent, programming also took place in the clinic (50, 12%), in the local community (45, 10%), in “other” settings such as a creative arts journal produced by students at a medical school (20, 5%), online (16, 4%), and in a mixture of these settings (54, 13%). Authors did not always specify the setting (61, 14%).

Programming was taught foremost by medical or health faculty (236, 56%). In these records, it was often not possible to determine from the information provided (e.g., author affiliations or body text) whether this group had training or background in the arts and/or humanities. Many records described collaboration among medical or health faculty and others such as humanities faculty, museum educators, and musicians (102, 24%).

Less often, the instructor was someone other than a medical or health faculty member (25, 6%), such as an artist or a hospice staff member. See Supplemental Digital Appendix 3 at http://links.lww.com/ACADMED/B94 for a complete instructor profile.

Programming was evaluated in about half the records (226, 53%). We did not analyze the data further to determine the type of evaluation used (e.g., learner satisfaction). Programming was not evaluated in 95 records (23%), and 103 records (24%) lacked sufficient information to determine whether the programming discussed was evaluated.

In most records, either the learners participating in the arts and/or humanities programming were not assessed (158, 37%) or the authors did not report whether the learners were assessed (140, 33%). Of the remaining records, 116 (27%) indicated that the learners were assessed, and 10 (3%) described multiple types of programming where learners were assessed for some types but not for others.

**Statistical analysis**

Explicit theoretical framing was significantly associated with both publication type and research type ($P < .001$). Books, book chapters, and dissertations/theses were more likely to integrate theory than journal articles. Descriptive records were less likely to ground the work in theory or in a theoretical framework, while conceptual records were more likely to do so.

**Stakeholder interviews**

We conducted 15 stakeholder interviews with 3 educators leading arts- and humanities-based curricula, 4 administrators supporting arts- and humanities-based curricula, 2 artists, 1 museum educator, 2 medical learners, and 3 arts and humanities scholars. Some participants exemplified more than 1 category.

Overall, the results of our descriptive and related statistical analyses aligned with participants’ experiences, notably in terms of the range of art forms and humanities subjects represented, the focus on the literary arts and on undergraduate medical education, and the sharp decline in arts- and humanities-based instruction during postgraduate education, for instance.

A dominant theme that emerged from the stakeholder interviews was their perception of a fundamental gap in the published literature. Participants noted that particular voices and experiences may be underrepresented in, or absent entirely from, the literature. This absence had 2 dimensions. First, the voices of those who teach and develop curricula but do not publish in traditional academic venues are missing from our scoping review findings. Participants suggested that these absent voices likely belong to artists, writers, museum educators, and humanities scholars and that, without them, educators may overlook the “role of creativity” in arts- and humanities-based work—what one docent-participant called “the process of a working artist and the kind of choice making and decision making” (P1) that artists enact. This observation aligns with and illuminates the potential implications of our finding that the literature was dominated by medical or health faculty writing about arts and humanities teaching.

Second, even within the voices present in the literature, stakeholders perceived that only selected stories are being told. One participant, an artist, suggested that a “publication bias” (P5) may hinder articles about the arts and humanities because program outcomes are challenging to measure quantitatively and journals may privilege such conventional outcomes. This potential bias affects not only what journals accept but also what authors submit for review. As the same participant explained:

> [W]hat do I think will get published? And, within my own power structure, what would be … appreciated or understood or comprehended? I don’t mean in … [the sense of a] “reward” …, but am I speaking the language that the powerful people around me can understand or value or fit [within] the competency compass. (P5)

Authors may thus prioritize scholarship that addresses aspects of arts and humanities curricula that can be objectively described and measured and therefore more readily integrated into ongoing scholarly conversations than more ineffable aspects of the field. A key factor influencing which studies are pursued and, from there, which stories are told may be that the majority of authors of the literature we reviewed were medical or health faculty members.
Discussion

There is a vast literature describing the uses of the arts and humanities to educate physicians across the continuum of medical education in Canada and the United States. Even after careful efforts to focus our review question and tighten our screening criteria, we still found 769 records to include in this scoping review. This represents a substantially larger sample than the mean of 117 records identified in a review of 494 scoping reviews published between 1999 and 2014.49 A positive implication of our sample size is that our synthesis results likely are an accurate reflection of the literature in this field.

However, a sample of this magnitude also has other implications, perhaps most importantly the trade-off between breadth and depth in our analysis. For example, we coded descriptions of arts and humanities curricula to determine which included an evaluation component, but we did not have the resources to further analyze those 226 records to determine the types of evaluation used. Similarly, we coded all records according to whether they were explicitly framed by theory or not, but we could not conduct further analysis of the 230 theoretically framed records to explore which theories underpin this literature. Such analyses remain for future scholarly efforts. We view the current synthesis as an authentic, albeit high level, reflection of the key patterns and gaps in this rich body of scholarship. We offer the following reflections on these patterns and gaps to inform current efforts to promote best practices in the use of the arts and humanities in medical education.

The diversity of the literature we reviewed represents a paradox. On the one hand, this diversity signals strength. The arts and humanities in medical education enjoy an abundance of forms and approaches, an infusion of knowledge from multiple disciplines, and an expansion of curricular activity across training levels and clinical specialties. On the other hand, this same diversity creates vulnerability. Although the phrase “arts and humanities” appears to signify a coherent field, the literature is compartmentalized into subfields, such as narrative medicine, graphic medicine, and visual thinking strategies, or clustered by art form or humanities subject, such as film and television, theater and drama, music, history, philosophy, and visual art. This phenomenon is not unique to the arts and humanities in medical education; fields divide in this way as scholars sub specialize their work. However, when the literature reads as a series of parallel conversations rather than a unified body of developing knowledge, it has implications for advancing a shared vision or national agenda for the arts and humanities in medical education.

Perhaps recognizing this paradox of diversity, Greene and Jones advocated a dual strategy:

- Efforts to define a role for medical humanities in medical education should … [both] continue to make the case for the shared contributions that all of the disciplines can make to medical education … [and] emphasize the valuable contributions of each specific discipline, in terms that medical educators can understand.52

Our analysis suggests that to achieve this dual strategy, those writing and publishing in the field need to improve the explicit and consistent use of terminology in the literature. We found that very few records used the terms “arts” or “humanities” intentionally: that is, these terms were rarely defined or located reflexively within a knowledge base. Eventually, we abandoned our efforts to infer what authors meant by their use of these ubiquitous terms.

This lack of explicit and consistent terminology persisted within subfields, as scholars have previously noted.51,52 Consider the example of narrative medicine. Some authors defined and/or referenced this term,53 some used it without explicit definition or reference,54 while others used terms similar to and suggestive of narrative medicine, such as narrative competence,55 narrative reflective practice,56 and narrative medical writing.57 Using the same terms in different ways, or using different terms to suggest the same thing, presents a barrier to building a coherent body of knowledge both within and across arts and humanities subfields.

Our analysis also revealed noteworthy patterns regarding arts and humanities curricular features. Most efforts were part of the undergraduate curriculum, with activities decreasing in frequency during postgraduate and continuing medical education. That arts and humanities curricula are particularly rare during continuing medical education suggests a worrisome disconnect.

Learners entering clinical training with undergraduate exposure to the arts and humanities may perceive that these disciplines are incompatible with clinical practice and their effects may be diminished if learners are no longer exposed to such programming during postgraduate and continuing education. In addition, most published curricula target medical trainees exclusively; rarely are the arts and humanities leveraged for interprofessional initiatives, according to the literature, although their relevance is broad and they could be used to advance shared concerns, such as the wellness of health care providers58 or the role of advocacy and social critique as part of health professional expertise.59

A persistent gap in evaluation and assessment is evident from our synthesis of the curricular features described across records. Only half of published programs appear to have been evaluated, and while we did not systematically assess the nature or strength of these evaluation studies, we noted a tendency toward reporting only learner satisfaction data. Learner assessment was less common. While many programs required learners to write, draw, act, reflect, or read, 70% of these programs either did not assess this participation or did not report having done so.

This pattern may reflect broader debates about how best—or even whether—to assess arts and humanities programming, given both its departure from medicine’s traditional assessment methods that quantify performance and the potential for the goals of learning to compete with those of assessment. For example, consider reflective writing. In the context of narrative medicine, Hermann wrote: “To rate or grade a piece of ‘reflective’ or creative writing, as is often done, is to distort the very idea of what writing in these contexts is ultimately for—discovery. How can you rate something as being more or less reflective?”61 Yet, medical educators commonly evaluate reflective writing in undergraduate portfolio courses. Faculty mentors provide regular formative
feedback to learners on their reflections, and at least 1 “theory-informed analytic rubric, demonstrating adequate interrater reliability, face validity, feasibility, and acceptability,” exists to evaluate learners’ reflective capacity based on their written reflections. As this example illustrates, assessment in arts and humanities curricula is complicated but not impossible.

Therefore, given that assessment is a powerful motivator in medical education, its absence in the scholarship is concerning. What educators could do most immediately is review and assess the products that learners create through arts- and humanities-based curricula. At minimum, educators could provide formative feedback on these products so that learners have an audience that responds to their work. Three previous reviews in the field noted a similar lack of evaluation and assessment data, fueling debate about the effectiveness of integrating the arts and humanities into medical education. We acknowledge the importance of considering effectiveness, and we encourage authors to include relevant evaluation data with their published curricula.

We also recognize that evaluation is a delicate topic, and educators need spaces to comment on their challenges in evaluating the arts and humanities in medical education contexts, given that conventional metrics could miss much of what these disciplines are positioned to do. However, we do not point to the lack of evaluation to rekindle these arguments. Rather, following Dennhardt and colleagues, we contend that the primary question at this time should be: What are the arts and humanities “trying to do” in medical education? Clarity on this fundamental point is necessary to support the right evaluation of the right outcomes.

Which brings us to 2 final, related insights. The first involves a gap that we can identify but not fully describe. As the stakeholders we interviewed pointed out, much arts and humanities activity in medical education never makes it into the published literature. Therefore, our study sample, despite its size, can only partially represent the work being done in the field. Precise characterization of what is missing is difficult. However, our findings suggest that, because medical faculty most often authored the published literature we reviewed, it is likely that a substantial part of what is missing is the voices of artists, docents, and other arts and humanities practitioners from outside medicine who contribute to these initiatives. With this realization comes another: The published literature regarding arts and humanities curricula represents a selection with a particular orientation. That most authors are medical faculty may account for the predominance of journal articles over other forms of scholarship. This finding also relates to the final gap we identified: the lack of theory in the literature.

Our statistical analysis demonstrated that journal articles in this field are significantly less likely than book chapters to explicitly integrate theory. While we cannot draw a causal relationship, we would nevertheless suggest that this overarching pattern of a predominance of journal articles (authored largely by medical faculty) and a paucity of theory shapes the scholarly conversation about the arts and humanities in medical education. This conversation is characterized by brief, episodic installments, privileging a medical orientation (over, say, an artistic orientation) and largely lacking a theoretical frame to weave the installments into a larger story that accumulates both over time and across subfields. A few subfields in the literature more regularly engaged theory as a strong component of scholarship (e.g., narrative medicine and theater), and smaller reviews as well as conceptual or qualitative papers offered theoretically informed frameworks for organizing the arts and humanities. This conversation is characterized by brief, episodic installments, privileging a medical orientation (over, say, an artistic orientation) and largely lacking a theoretical frame to weave the installments into a larger story that accumulates both over time and across subfields. A few subfields in the literature more regularly engaged theory as a strong component of scholarship (e.g., narrative medicine and theater), and smaller reviews as well as conceptual or qualitative papers offered theoretically informed frameworks for organizing the arts and humanities.

However, missing from the literature is an overarching theory of practice that engages with and is relevant to all of the diverse subfields that make up the arts and humanities in medical education. We address this gap directly in a separate publication.

Limitations

This scoping review is shaped by a number of design decisions. Some are derived from the review’s status as scholarly work commissioned by a national association (AAMC). Regular discussions with this knowledge user (and funder) were part of the review process, and fundamental decisions, such as the wording of the research question, the 12-month timeline, and the resulting rapid review procedures, were strongly influenced by the AAMC’s aims and requirements. The databases we searched were restricted to those that could accommodate the search strategy; some relevant databases such as JSTOR could not accommodate the search strategy’s multiple strings and multiple keywords within each string so they were excluded from the search. The scale of the project combined with the short timeline required the use of strategies from rapid review methodology to ensure feasibility and completion of the work. Consequently, the search was limited by time, English language, and geography, and we neither included gray literature nor hand-searched bibliographies. Such decisions, notwithstanding the study sample size, constrain the conclusions we can draw regarding the field as a whole.

Because a scoping review does not engage in an evaluation of the quality of the included records, we relied on what the authors reported to determine categorizations, such as review type, research method, etc. To the extent that these authors may have used terms inconsistently or labeled their methodology unclearly, our descriptive categorization of the records may be incorrect. There is subjectivity inherent in applying any code to extract data, and although we have presented descriptive statistics in our summary of this coding, we consider these a sketch of the published literature rather than a precise measure of it. That the stakeholders we interviewed corroborated our descriptive findings lends credibility to our analyses and suggests that the patterns we identified are a reasonable starting point for future research efforts.

We also engaged in informal (nonaudiorecorded) member checking through our presentations to a variety of knowledge user groups; discussions in these settings enriched our interpretive work at multiple points in the analytical process. However, because some stakeholders we interviewed had participated in these public events, their access to our preliminary analyses was different than other participants and may have shaped their responses in the interviews.

Finally, the statistical analysis was exploratory, so we did not adjust for
potential confounding factors, and there were deviations from the assumptions of linear regression, in particular for the normality of the data. However, we explored the use of Poisson negative binomial regression, and the results matched those of the linear regression. The addition of a constant value was required to execute log transformations, which has been shown to slightly reduce variance; however, we used a $P < .01$ threshold to reduce the risk of type I error.

**Future research**

Future work should include a review of gray literature and a study of the literature beyond Canada and the United States. We excluded 550 international records at the level of full-text screening, with additional records excluded at the title and abstract review. While we used the expert knowledge of our review team to ensure that we captured foundational pieces from outside Canada and the United States, there is more to learn about the uses of the arts and humanities in medical education by exploring international literature.

Furthermore, we analyzed whether arts and humanities programming was evaluated but not how it was evaluated for the 226 records with evaluation data. Future research should explore evaluation in more depth to learn what educators are evaluating and how they are going about it. While a previous review explored quantitative evaluation methods, further study is needed to explore whether and how educators are evaluating learners in less conventional ways. Finally, our data set offers the opportunity for integrative analyses of the various functions that arts and humanities curricula can serve across the medical education continuum.

**Conclusions**

The findings of this scoping review demonstrate at a high level the extent, range, and nature of the published scholarship on the uses of the arts and humanities across the medical education continuum in Canada and the United States. While the uses of the arts and humanities in this published literature are rich and diverse, the knowledge arising from these activities is relatively impoverished due to compartmentalization, a lack of theory, and missing perspectives. More effort is needed in 3 areas: to build knowledge across subfields (e.g., through consistent and shared language), to substantively engage theory (e.g., through the development of an overarching theory of practice for the field), and to include the voices of artist-practitioners alongside medical educators (e.g., through collaborative research and publication with those artist-practitioners involved in curriculum development and delivery). These key patterns and gaps should inform future efforts to promote and study the uses of the arts and humanities in medical education.

**Acknowledgments:** This scoping review was initially conceptualized by a larger team: Nancy Adams, Tavis Apramian, Shannon Arnfield, Claire de Boer, Esther Dell, Maryam Golafshani, Paul Haidet, Lorelei Lingard, Tracy Moniz, Javed Sukherra, and Rebecca Volpe. The authors wish to thank Tavis Apramian, Western University; for contributions to study selection; Shannon Arnfield, Western University; for contributions to data interpretation; Niklas Bobrovitz, Massey College and University of Toronto, for conducting the statistical analysis; Francesco Colosimo, Western University; for contributions to database organization and study selection; Esther Dell, Penn State College of Medicine, for peer reviewing the scoping review search strategy; and Markus Gulliat, University of Toronto, for data visualization support.

**Funding/Support:** This research was funded in part by the Association of American Medical Colleges.

**Other disclosures:** None reported.

**Ethical approval:** Research ethics clearance was granted from the University Research Ethics Board at Mount Saint Vincent University on June 27, 2019 (file #2019-015). This research received exemption from the Institutional Review Board at Penn State College of Medicine on November 11, 2019 (#STUDY00013567).

**Disclaimers:** The views expressed herein are those of the authors and do not necessarily reflect the position or policy of the Association of American Medical Colleges.

**Previous presentations:** Preliminary results from the scoping review were presented at 2 meetings (September 13, 2019, and February 4, 2020) of The Fundamental Role of the Arts and Humanities in Medical Education Integration Committee, Association of American Medical Colleges, Washington, DC, and at the Learn Serve Lead annual meeting of the Association of American Medical Colleges on November 10, 2019 in Phoenix, Arizona, as well as published as an oral abstract in the Canadian Medical Education Journal, Canadian Conference on Medical Education, on April 18, 2020.

**T. Moniz** is associate professor, Department of Communication Studies, Mount Saint Vincent University, Halifax, Nova Scotia, Canada; ORCID: https://orcid.org/0000-0002-5078-4611.

**M. Golafshani** is a second-year medical student, Faculty of Medicine, University of Toronto, Toronto, Canada.

**C.M. Gaspar** is a PhD candidate, Faculty of Health, Dalhousie University, Halifax, Nova Scotia, Canada; ORCID: https://orcid.org/0000-0003-0412-5495.

**N.E. Adams** is associate librarian and assistant dean of foundational sciences, Penn State College of Medicine, Hershey, Pennsylvania; ORCID: https://orcid.org/0000-0003-0534-6716.

**P. Haidet** is director of medical education research, Woodward Center for Excellence in Health Sciences Education, Penn State College of Medicine, Hershey, Pennsylvania.

**J. Sukherra** is associate professor, Departments of Psychiatry and Paediatrics, and scientist, Centre for Education Research and Innovation, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada; ORCID: https://orcid.org/0000-0001-8146-4947.

**R.L. Volpe** is associate professor and vice chair for education, Department of Humanities, Penn State College of Medicine, and director, Clinical Ethics Consultation Service, Penn State Health Milton S. Hershey Medical Center, Hershey, Pennsylvania; ORCID: https://orcid.org/0000-0003-3406-9498.

**C. de Boer** is director, The Doctors Kienle Center for Humanistic Medicine, and founding director, Center Stage Arts in Health, Penn State College of Medicine and Penn State Health Milton S. Hershey Medical Center, Hershey, Pennsylvania. She is president, National Organization for Arts in Health, San Diego, California.

**L. Lingard** is professor, Department of Medicine, and scientist, Centre for Education Research and Innovation, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada.

**References**


8. Haidet P, Jarecke J, Adams NE, et al. A guiding framework to maximise the power of...


