**APPENDIX 2**

**Research Protocol**

**Background**

Telehealth is known to help patients overcome two barriers they face, when seeking health care: distance and time. Additionally, telehealth has been found to deliver a range of benefits, including improved access to healthcare, more timely interactions between providers and patients, improved continuity of care, better health outcomes, higher patient & provider satisfaction, and more efficient use time, to name a few . Proponents of telehealth have argued that it has the potential to transform healthcare delivery by reducing costs and increasing quality and patient satisfaction.

In the United States, although telehealth was a ‘hot topic’ before COVID-19, its use was far from widespread, and was limited to certain medical specialties. For example, a pre-pandemic US-based weighted survey study on use of telemedicine by individual physicians in their practice (published in *Health Affairs* in 2018), found wide variations in the use of telemedicine (for physician-patient interactions) by medical specialty, ranging from a low of 6.1% in the internal medicine subspecialty of *Allergy-Immunology* to a high of 39.5% in *Radiology*. Other “lower-using” specialties identified by the study were *Gastroenterology* (7.9%), OB/GYN (9.3%), *General Surgery* (9.7%), and *Family Medicine* (11.8%). On the other hand, other “higher-using” specialties included *Psychiatry* (27.8%), the internal medicine subspecialty of *Cardiology* (24.1%), *Pathology* (23%) and *Emergency Medicine* (22.3%). The results of this study have already been widely cited in the general telehealth literature and corroborated by evidence and position statements in the specialty-specific telehealth literature as well.

The wide variation in telehealth use across medical specialties is intriguing, since the U.S. lacks a nationwide standardized set of telehealth reimbursement policies, which in turn, has historically posed a barrier to telehealth adoption in all specialties. Despite these macro (policy-level) constraints experienced by all medical specialties, some specialties have been able to normalize telehealth to mainstream practice (e.g., *Psychiatry*, *Cardiology*); while others are just getting started during the COVID-19 pandemic (e.g., *Allergy-Immunology*), despite potential for all stakeholders within the specialty (e.g., patients, providers) to benefit from telehealth adoption

With the arrival of COVID-19, the healthcare industry has witnessed a massive acceleration in the use of telehealth, not only because of physical (social) distancing requirements, but also due to the temporary removal of policy & regulatory barriers to telehealth access. Nevertheless, the questions that do not yet have definitive answers are: Is telehealth here to stay? Will providers find it to be an effective method for providing care? Will patients prefer it over traditional in-person visits? Will policy barriers to telehealth access be permanently lifted? Although there is much uncertainty about the future of telehealth in the post-pandemic era, one point of consensus is that the elimination of policy barriers by itself, would not suffice to enable widespread, sustainable use of telehealth services. Instead, the telehealth literature has emphasized the need for providers and organizations to make concerted efforts to design and implement telehealth services for sustainable use. Not surprisingly therefore, a growing stream of pandemic literature has emphasized the importance of addressing the challenge of wide variations in telehealth use, for ensuring widespread sustainability in the post-pandemic future.

***Research Questions***

To be able to identify strategies for successful and sustainable telehealth use across all specialties in a post-pandemic era, it would be important to understand the “specialty-level factors” that historically influenced telehealth use among both “lower-using” specialties and “higher-using” specialties. To clarify, the term “telehealth use” in this paper, refers to adoption and utilization of telehealth services by providers (and patients) within a given medical specialty. Given the above background information, the two research questions are:

1. What specialty-level factors historically influenced telehealth use in six medical specialties, including three “higher-using” specialties (*Cardiology, Psychiatry, Radiology*) and three “lower-using” specialties (*Allergy-Immunology, Gastroenterology, Family Medicine*), in the U.S.?
2. Which factors (barriers or facilitators) are associated with relatively lower telehealth use in some specialties and relatively higher telehealth use in other specialties?

By understanding the barriers to telehealth use among “lower-using” specialties and facilitators to telehealth use among “higher-using” specialties, we seek to identify strategies for reducing variation in telehealth use across all medical specialties, and discern implications for ensuring widespread, sustainable use of telehealth services in a post-pandemic environment.

***Purpose & Aims***

In keeping with the research questions, the purpose of this paper is to conduct a systematic review and narrative synthesis of the literature to identify the key factors historically influencing telehealth use in six medical specialties (selected to represent a wide variation in historical telehealth use), including three “lower-using” specialties (Allergy*-Immunology, Gastroenterology,* and *Family Medicine*) and three “higher-using” specialties (*Cardiology, Psychiatry, Radiology*). The specific aims of the review are to:

1. Describe the factors (including barriers and facilitators) historically influencing telehealth use in six medical specialties in the U.S., including three “lower-using” specialties and three “higher-using” specialties.
2. Identify strategies for reducing variation in telehealth use across medical specialties in a post-pandemic era.
3. Discuss implications for ensuring widespread, sustainable telehealth use across medical specialties, in a post-pandemic era.

**Methodology**

Design

We conducted a systematic review and narrative synthesis of the literature, using the guidance developed by Popay et al (2006), to examine factors historically influencing telehealth use across medical specialties. Narrative synthesis refers to ‘*an approach to the systematic review and synthesis of findings from multiple sources and relies primarily on the use of words and text to summarize and explain the findings of the synthesis*.’ It is used when statistical meta-analysis or another specialized form of synthesis (e.g., meta-ethnography for qualitative studies) is not feasible, due to high degrees of breadth and heterogeneity in the literature. Given our broad topic of interest, our literature of interest was also broad, including peer-reviewed original studies, reviews, and specialty-society workgroup reports & position statements. Our focus was to review all forms of peer-reviewed literature that helped to understand the factors that historically influenced telehealth use in each specialty.

Data sources

Article searches were conducted on the PUBMED database. **Table 1** summarizes our search strategy. As indicated in the table, our initial database search entailed using key search terms (“Telehealth,” “<Specialty Name>,” “Barriers,” “Facilitators”) and closely-related terms informed by our conceptual framework (**Figure 1**). PRISMA guidelines were used to inform the process of article selection based on study inclusion criteria (outlined below). The PRISMA checklist is included in **Appendix 1**.

**Table 1: Database Search Strategy**

|  |  |
| --- | --- |
| **Key search terms** | **Closely-related terms** |
| Telehealth | Telemedicine |
| Barriers, Facilitators | Influencing factors, including *macro* (policy-level) factors, *meso* (organizational-level) factors; and *micro* (individual-level) factors. |
| Specialty Name: *Allergy-Immunology, Gastroenterology, Family Medicine, Cardiology, Psychiatry, Radiology*. | Specialty Name: *Asthma,* *Allergy, Hepatology, Primary Care, Cardiovascular, Behavioral Health, Mental Health, Imaging.* |

Developing a narrative synthesis

We followed the steps delineated by Popay et al (2006), namely (1) Identifying a conceptual framework, (2) developing a preliminary synthesis (which included identifying inclusion & exclusion criteria), (3) exploring themes in the data, and (4) assessing robustness of final synthesis [28]. We used the *macro-meso-micro* three-layer framework (available since the beginnings of social science research), to guide our review. This framework (**Figure 1**), represents three layers of factors influencing behavior (e.g., telehealth use in a medical specialty). The *macro* layer represents factors at the societal level (e.g., *policy* or *structural* factors); the *meso* layer represents factors at a group level (e.g., *organizational* factors in a given medical specialty), and the *micro* layer represents factors at the individual level (e.g., *provider-* or *patient-level* factors).The framework aims to capture three layers of factors as well as the interrelationships among those layers in influencing behavior. It also highlights the role of context in shaping interrelationships across the three layers, to ultimately influence behavior.

Conceptual Framework

As discussed in the Background, the telehealth literature points to key macro-layer factors (e.g., policy-level factors like telehealth reimbursement) in influencing telehealth use across all specialties. Logically therefore, it would be essential to also identify key meso-layer factors (e.g., organizational or specialty-level factors) and key micro-layer factors (e.g., individual provider and patient-level factors) influencing telehealth use, to arrive at a comprehensive set of factors historically influencing telehealth use in higher- and lower-using specialties.

Ly et al (2017) used the *macro-meso-micro* framework to characterize the factors influencing telemedicine use in Senegal. In their framework, the *macro* layer included political, financial, legal, and ethical, factors; the *meso* layer included organizational and technical factors; while the *micro* layer included individual intention to use telemedicine. This framework provided a starting point for a preliminary synthesis of the literature, which in turn enabled further refinement of the framework in regard to specialty-level factors of relevance to this review. The final conceptual framework (**Figure 1**) used to inform the narrative synthesis comprised of a total of 12 factors across the three layers.

* The *macro* layer included 3 structural factors (policy, legal, and other structural factors).
* The *meso* layer included 7 organizational or specialty-level factors (historical orientation to telehealth, hospital-organizational, professional-society, telehealth outcomes-research, cultural, treatment, and technological factors).
* The *micro* layer included 2 individual-level factors (provider- and patient-level factors).

**Figure 1: Conceptual Framework**



Preliminary Synthesis (Inclusion & Exclusion Criteria)

The initial version of the conceptual framework (adapted from Ly et al, 2017) was used to screen articles and conduct a preliminary synthesis of articles resulting from the PUBMED database search. As indicated in **Figure 1**, this process helped to expand upon the listing of meso (specialty-level) factors, from 2 factors (in the initial framework), to the following 7 factors in the final conceptual framework, which provided the basis for narrative synthesis.

1. Specialty-level historical rationale for telehealth
2. Specialty-level hospital-organizational factors
3. Specialty-level professional-society factors
4. Specialty-level telehealth implementation & outcomes-research factors
5. Specialty-level cultural factors
6. Specialty-level treatment factors
7. Specialty-level technological factors

In addition to refining the conceptual framework, the preliminary synthesis helped to determine inclusion/exclusion criteria for the narrative synthesis. At the outset, abstracts of articles resulting from the PUBMED database search were screened by the three authors (reviewers), after removal of duplicates, to assess eligibility. As mentioned earlier, all forms of peer-reviewed literature that helped to address the two research questions (e.g., original studies, reviews, and specialty-society workgroup reports & position statements), were included. Since the interest was in understanding factors *historically* influencing telehealth use across medical specialties, the search was not restricted by date. On the other hand, in keeping with the scope of the review (within the U.S.), articles that did not originate in the U.S., were excluded. Additionally, articles with a narrow focus were excluded, for example, if they described specific results of telehealth-related outcomes research, e.g., for special populations or specific modalities. To clarify, since the primary research question was broad—to gain an understanding of factors historically influencing telehealth use across specialties, it was more relevant to understand *how many* such studies (telehealth-related clinical trials) existed in a given specialty over a 10-year period preceding COVID-19 and to understand the state of telehealth-related outcomes research in that specialty, rather than to know the specific results of such studies. The above exclusion criterion helped to add ‘research factors’ to our conceptual framework at the meso (specialty) level. Lastly, articles were excluded if they did not shed light on either research question, or if they duplicated the insights already gained.

Moving beyond refinement of the conceptual framework and identification of inclusion & exclusion criteria, the preliminary synthesis entailed completion of two steps: 1) extracting key article characteristics to develop a descriptive summary of articles reviewed; and 2) generating a textual summary of the articles (using the conceptual framework), to enable exploration of themes to address the two research questions. The former helped to develop a template for describing articles reviewed in the final narrative synthesis while the latter helped to develop *rules* for: 1) exploration of themes among articles within and across specialties and 2) assessment of robustness of the final narrative synthesis.

Data Extraction and Exploration of Themes

As mentioned earlier, preliminary synthesis involved completing two steps: 1) extracting article characteristics to develop a descriptive summary of articles reviewed; and 2) generating a textual summary of the articles for exploration of themes.

*Extraction of article (descriptive) characteristics*: The following information was extracted from articles reviewed per specialty (1) publication citation; 2) type of article; 3) layer of emphasis (macro, meso, or micro); 4) research question addressed; 5) descriptive summary of article. This provided a template for descriptive summaries of articles reviewed for the final narrative synthesis. Extracted data was reviewed by all authors for disparities or inconsistencies. As a next step, a textual summary of articles was generated, to enable exploration of themes in regard to factors historically influencing telehealth use, both within and across specialties. This process in turn, helped to develop rules for assessing the robustness of final narrative synthesis.

*Generation of textual summary for thematic analysis:* Next, the conceptual framework was used as a basis for exploring themes among articles within and across specialties, to address the two research questions. Articles within each specialty were reviewed for information pertaining to any of the 12 factors encompassed in the macro, meso, and micro layers. This process helped to generate a textual summary of articles within each specialty (for each of the 12 factors), which in turn, enabled exploration of themes within and across specialties. **For example**, with respect to the “specialty-level historical orientation towards telehealth” in the meso layer, several articles within the *Allergy-Immunology* specialty (a “lower-using” specialty) indicated that the historical rationale for telemedicine use in this specialty has been to improve “access to care” for poor, elderly, and disabled. This in turn, conveyed a limited view of telehealth as being applicable only to a certain segment of population to address access-to-care issues. By comparison, several articles in *Cardiology* (a “higher-using” specialty), indicated that telehealth has historically been used in this specialty to improve quality of care, including improving patient outcomes & experience through disease management and reduction of hospitalizations for heart failure. Additionally, the *Cardiology* literature revealed that over the past two decades, hospital organizations have favored telehealth use in this specialty, as a result of federal initiatives towards public reporting and pay-for-quality, spurred by legislation such as the Affordable Care Act (ACA) of 2009 and the Medicare Access and CHIP Reauthorization Act (MACRA) of 2015, which created strong incentives for using telehealth to improve patient outcomes (e.g., reducing readmissions for heart failure). The latter in turn, served a dual purpose of improving quality and reducing costs, thereby creating a business case for hospitals to use telehealth in *Cardiology*. In summary, the literature revealed that providers and organizations in the *Cardiology* specialty recognized early on, that telehealth had the potential to improve quality, improve patient experience (promote patient satisfaction), and reduce costs.

The above is an example of how the themes of “access to care,” “quality improvement,” and “cost reduction” were identified under the meso layer factor of “specialty-level historical orientation to telehealth.” The above example also serves to illustrate how themes were identified based on exploration of relationships among articles within and across specialties, for each factor in the conceptual framework. For example, the theme of “access to care” was repeated in three or more articles within the *Allergy-Immunology* specialty, while the themes of “quality improvement” and “cost reduction” were repeated in three or more articles within the Cardiology specialty. This in turn, helped to consolidate our understanding of differences in “specialty-level historical orientation to telehealth” across two subspecialties of Internal Medicine, i.e., *Allergy-Immunology* (a “lower-using” specialty) and *Cardiology* (a “higher-using” specialty). Importantly, this process helped to develop rules for assessing the robustness of final narrative synthesis.

Assessing the Robustness of Final Narrative Synthesis

The textual summary of articles was maintained in a Microsoft Excel workbook. Each author maintained a separate workbook. All three authors independently reviewed included articles (in all six specialties) for final synthesis. Each individual workbook included six worksheets dedicated to each specialty of interest. In keeping with the conceptual framework, each worksheet was organized by layer and factor (for a total of 12 factors under the three layers). Each worksheet contained a textual summary for every factor in the cell block adjacent to the factor name. Additional data elements in each worksheet included, number of articles informing each textual summary; lead author and year of articles informing the textual summary; number of themes; theme-headings identified in the textual summary; and author (reviewer) initials. As indicated earlier, a theme was confirmed if it was repeated in three or more articles within each specialty.

The robustness of final synthesis was assessed in three stages. In Stage 1, all three authors (*Author-1, Author-2,* and *Author-3*) independently reviewed all articles in each of the six specialties to develop textual summaries for each factor in each specialty worksheet and complete the remainder of each worksheet. In Stage 2, the three authors met several times to discuss their respective textual summaries for every factor within each specialty. The purpose of these discussions was to identify the number of themes that were overlapping and the number of themes that were distinct for each textual summary (cell block) in each specialty worksheet, across each of the three author-pairs (*Author-1 & Author-2; Author-1 & Author-3; Author-2 & Author-3*). During the course of these discussions, all overlapping themes were acknowledged, and distinct themes were verified through re-review of relevant articles. This process was repeated for all textual summaries in the six worksheets. **For example,** in the context 0f “specialty-level historical orientation to telehealth” (discussed earlier) in the meso layer, if *Author-1* & *Author-2* had identified the sub-theme of “access to care” the latter would be considered an overlapping theme, on the other hand, if *Author-3* had identified “improving patient experience” (not identified by first two authors), the latter would be considered a distinct theme. In Stage 3, based on all Stage 2 discussions, the first author developed an integrated Excel workbook to include the worksheets from other two authors for every specialty. Additionally, the first author worksheet for every specialty was updated to include three columns (one for each author-pair), to indicate the number of overlapping and distinct themes relevant to each textual summary.

All aforementioned rules developed for data extraction, exploration of themes among articles within and across specialties, and assessment of robustness of synthesis, were applied to the final narrative synthesis of articles reviewed across six medical specialties. The integrated Excel workbook (described above) constituted the raw dataset for final narrative synthesis in this review. It contained 18 worksheets (3 for each of the 6 specialties), and is included in supplementary materials as **Appendix 3**.