Routledge
Taylor & Francis Group

Health Communication

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hhth20

Comparisons of Communication in Medical Face-To-Face and Teleconsultations: A Systematic Review and Narrative Synthesis

Joseph Ford & Markus Reuber

To cite this article: Joseph Ford & Markus Reuber (2023): Comparisons of Communication in Medical Face-To-Face and Teleconsultations: A Systematic Review and Narrative Synthesis, Health Communication, DOI: 10.1080/10410236.2023.2201733

To link to this article: https://doi.org/10.1080/10410236.2023.2201733

Published online: 24 Apr 2023.

HEALTH COMMUNICATION
https://doi.org/10.1080/10410236.2023.2201733

Routledge
Taylor & Francis Group

Comparisons of Communication in Medical Face-To-Face and Teleconsultations: A Systematic Review and Narrative Synthesis

Joseph Ford and Markus Reuber

Department of Neuroscience, University of Sheffield, Broomhall, Sheffield, UK

ABSTRACT

The COVID-19 pandemic has brought telemedicine into mainstream medical practice (although questions remain over its role in a post-pandemic world). Research suggests that most patients and providers are satisfied with the flexibility and convenience of teleconsultations. However, there is continuing uncertainty about whether this shift has a clinically relevant impact on the quality of doctor-patient interaction. We conducted a systematic search of studies comparing communication in medical face-to-face consultations and teleconsultations. We included only studies which examined communication directly using recordings, excluding studies which used questionnaires or interviews. Studies were appraised using modified versions of the Critical Appraisal Skills Programme (CASP) checklists. Our search yielded 25,348 records, of which 22 were included in the final review. These studies were conducted in various medical specialties. Methodologies included approaches based on quantified communication behaviors using coding systems and qualitative studies using microanalytic methods. Except for duration (where there was evidence of face-to-face consultations being longer), no differences between the two modes of communication were consistently identified. In the aggregate, however, statistically significant differences were more prominent in primary care and more likely to favor face-to-face consultations. Qualitative studies also highlighted differences in how communication behaviors were manifested in each modality. Because much of the examined research was conducted in selected or non-routine settings, its applicability to the less selective use of telemedicine during and after the pandemic is limited.

Introduction

The COVID-19 pandemic has prompted a dramatic increase in medical consultations carried out online or over the telephone rather than face-to-face (although questions remain over the role of telemedicine in a post-pandemic world). Research has highlighted the convenience and accessibility of teleconsultations (Gomez et al., 2021; Harrison et al., 2006), suggesting high levels of patient (Agha, Schapira, et al., 2009; Nakornchai et al., 2021; Nguyen et al., 2020) and provider satisfaction (with caveats – see Bulik, 2008; Courtney et al., 2021; Nguyen et al., 2020). There are concerns, however, that teleconsulting may impede key aspects of medical practice such as examination (Courtney et al., 2021; Gomez et al., 2021) and the doctor-patient relationship (Gomez et al., 2021; Miller, 2010). Technical problems may impede communication, and the reliance on technology could lead to increases in health disparities between people with access to high quality communication equipment and people without (Bailey et al., 2015).

Taken together, this evidence presents a mixed picture of teleconsulting. It is noteworthy that the positive and negative aspects listed above are not incompatible, with the former highlighting the practical benefits of teleconsultations and the latter focusing on the impact that they may have on the consultation process (e.g. diagnosis) and, potentially, treatment outcomes. As Shaw et al. (2020) note, "although the evidence on video consultations indicates that they are feasible, safe, and effective in health care, the evidence on interaction in such consultations is limited."

These aspects of teleconsultations can be studied using interaction coding systems (e.g. the Roter Interaction Analysis System or RIAS (Roter & Larson, 2002)) and qualitative methods such as Conversation Analysis (CA), which seeks to explicate the ways in which doctors and patients carry out clinical tasks (e.g. diagnosis; Peräkylä, 1998; Reuber et al., 2015). These methods have been widely applied to both traditional face-to-face medical communication (Beach, 2012; Heritage & Maynard, 2006; Pires & Cavaco, 2014) and, to a lesser (Miller, 2003, 2010) extent, to teleconsultations (Miller & Nelson, 2005; Nelson et al., 2010).

These methods are also appropriate for examining whether and how face-to-face and teleconsultations *differ* from each other. Answering this question does, of course, require a comparative approach. Comparative studies will need to include methods with a microanalytic focus, as earlier research has shown how even subtle differences in communication, down to individual word selection, can affect how a consultation unfolds (Heritage et al., 2007). Once differences have been identified by the detailed analysis of clinical interactions, other (perhaps mixed) methods would need to be employed to examine to what extent such differences influence clinically relevant outcomes.

CONTACT Joseph Ford j.w.ford@sheffield.ac.uk Department of Neuroscience, University of Sheffield, 385a Glossop Rd, Broomhall, Sheffield S10 2HQ, UK

The aim of this article is to review studies that have carried out empirical comparisons of doctor-patient communication and to identify what differences, if any, they have identified between face-to-face and teleconsultations.

Materials and methods

Databases and search terms

As summarized in Table 1, we carried out a systematic search of eight databases in the fields of medicine and social sciences. Note that the search terms differed slightly for journals searched using ProQuest because ProQuest's truncation symbol (*) only allowed for five additional characters. Where possible, we also used MeSH/thesaurus terms (though, as can be seen in Table 1, this did not always yield additional records). There were no limitations on publication date, meaning that all studies meeting the criteria and published up until the date of the search (29/3/21) were eligible for inclusion.

Paper screening

Figure 1 shows our paper screening process. Four criteria were established to screen papers. First, studies had to be comparative between face-to-face and teleconsultations. Second, the comparison had to be based on empirical data rather than, for example, anecdotal experiences. Third, the empirical data had to comprise recordings of face-to-face and telemedicine encounters. Fourth, the focus of the study had to be on patient-provider communication, thus excluding studies where recordings of consultations were used to measure and compare a diagnostic variable (e.g. a patient's speech fluency in speech-and-language therapy).

Data extraction

Data about each study were extracted and compiled. In cases where other methods (e.g. interviews) were used alongside the interactional analysis, only the data relating to the interactional aspect was extracted. See Table 2 for a summary of the studies.

Quality appraisal

Studies were appraised using modified forms of the Critical Appraisal Skills Programme (CASP) checklists. The checklists used were the CASP Randomised Control Trials Checklist (Critical Appraisal Skills Programme, 2020), the CASP Qualitative Studies Checklist (Critical Appraisal Skills Programme, 2018b), and the CASP Case Control Study Checklist (Critical Appraisal Skills Programme, 2018a). These checklists were modified to remove and reword certain questions, as well as adding a series of questions specific to the concerns of this review.

Analytic approach

This review was conceptualized as a broad exploration of both quantitative and qualitative studies. Researchers have noted three designs for mixed-methods reviews: segregated, integrated, and contingent (Sandelowski et al., 2006). We applied an integrated design, the defining feature of which is that "the methodological differences between qualitative and quantitative studies are minimized as both kinds of studies are viewed as producing findings that can be readily transformed into each other" (Sandelowski et al., 2006, p. 8).

In this case, both quantitative coding studies and qualitative interaction analyses could address our research question, with

Table 1. Databases, search fields, search exclusions, MeSh/equivalent terms, and search terms used when searching literature.

Database	Fields searched	Excluded from results	MeSH/ equivalent terms	Search terms
Linguistics Abstracts Online	EBSCO default (abstract, title, keywords)	Magazines (1)	Made no difference	(communic* OR interact* OR conversat* OR talk* OR empath* OR rapport OR discourse OR discursive OR relation*) AND (telemedic* OR telehealth* OR teleconsult* OR "telephone consult*" or "phone consult*" OR "remote consult*" OR "virtual consult*" OR "video consult*")
CINAHL	EBSCO default (abstract, title, keywords)	None	Found five additional records	
Psycinfo	OVID's '.mp' field code (abstract, title, keywords, heading word, table of contents, key concepts, original title, tests and measures, and MeSH terms)	None	Made no difference	
Scopus	Scopus's TITLE-ABS-KEY search	Reviews (2445), Notes (472), Editorials (427), Letters (379), Short Surveys (194), Conference Reviews (141), Errata (5), Data Papers (1), Retracted papers (1)	N/A	
MEDLINE	EBSCO default search (abstract, title, and keywords)	Magazines (75) and Guidelines 920)	N/A	
Applied Social Sciences Index and Abstracts	ProQuest's 'noft' field code (everything but full text)	None	N/A	(communic* OR interact* OR conversat* OR talk* OR empath* OR rapport OR discourse OR discursive OR relation*) AND (telemedic* OR telehealth* OR teleconsult* OR "telephone consult*" or "phone consult*" OR "remote consult*" OR "virtual consult*" OR "video consult*" OR "teleconsultation" OR "telephone consultation*" OR "phone consultation*" OR "remote consultation*" OR "virtual consultation*" OR "video consultation*")
Sociological Abstracts	ProQuest's 'noft' field (everything but full text)	None	N/A	
Social Science Database	ProQuest's 'noft' field (everything but full text)	None	N/A	

Table 2. Papers reviewed.

Paper	Setting	Country	Type(s) of telemedicine	No. of participants	Data	Approach
Quantitative studies						
Agha, Roter et al. (2009)	Pulmonary consultations in veterans' hospitals	USA	Video	19 patients (11 video, 8 face-to-face)	Video recordings of consultations	Coding: • Roter Interaction Analysis System (RIAS) Other: • Measured duration and verbal dominance
Agha et al. (2010) [Conference abstract]	Veterans' hospital pulmonary, endocrine, and rheumatology clinics	USA	Unsure	221 patients (111 video, 110 face-to-face)	Video recordings of consultations	Coding: • Roter Interaction Analysis System (RIAS)Other: • Measured duration and verbal dominance
Ball et al. (1995)	Acute psychiatric unit	UK	Telephone, hands-free telephone, and video	6 patients (all patients took part in all conditions)	Video and audio recordings of consultations	Coding: • Stiles Verbal Response Modes (VRM) for verbal behavior • Video Interactive • Techniques for Automated Scoring System (VITAS) for non-verbal behavior
Day and Schneider (2002)	Psychotherapy	USA	Audio and video	80 clients (27 face-to-face, 26 video, 27 telephone) Each client received five sessions each	Video recordings	Coding: • Vanderbilt Psychotherapy Process Scale
Demiris et al. (2005)	Dermatology	USA	Video	94 patients (40 face-to-face, 54 telemedicine)	Audio recordings	Coding: • Protocol inspired by the Davis Observation CodeOther: • Measured duration and verbal dominance
Edison et al. (2013)	Dermatology	USA	Video	91 patients (41 face-to-face, 50 telemedicine)	Audio recordings	Coding: • Protocol inspired by the Davis Observation Code Other: • Measured duration and amount of talk from each party
Frueh et al. (2007)	Group therapy	USA	Video	38 clients (21 face-to-face, 17 video) Clients participated in 17 sessions total	Audio recordings	Coding: • Social and emotional rehabilitation-therapist adherence and competency protocol (SER-TACP)
Hammersley et al. (2019)	Primary care follow-up consultations	UK	Video and telephone	149 patients (51 face-to-face, 53 telephone, 45 video)	Audio recordings	Coding: • Royal College of General Practitioners (RCGP) quality indicators • RIAS Other: • Measured duration
Innes et al. (2006)	Primary care	UK (this study), USA (existing comparison study)	Telephone	43 telephone patients Comparisons were made with findings from earlier study of face-to-face consultations involving 277 patients	Audio recordings	Coding: • RIAS
Laflamme et al. (2005)	Nursing home	USA	Video	41 nursing home residents (12 face-to-face, 29 video)	Video recordings	Coding: • Bespoke 31-item instrument • Measured duration
Liu et al. (2007)	Primary care (internal medicine)	Japan	Video	20 patients All patients took part in both conditions, with 10 having face-to-face first and 10 having video first (total of 40 consultations)	Video recordings	Coding: • Bespoke system based on Brink-Muinen et al.'s (2002) work on verbal behavior (included duration)

(*Continued*)

Table 2. (Continued).

Paper	Setting	Country	Type(s) of telemedicine	No. of participants	Data	Approach
McKinstry et al. (2010)	Primary care	UK	Telephone	105 patients (59 face-to-face, 46 telephone)	Audio recordings	Coding: • RIAS • RCGP criteria • Observing Patient Involvement in Decision Making (OPTION) measure Other: • Measured duration and number of problems
McKinstry et al. (2011)	Primary care	UK	Telephone	144 (94 face-to-face, 50 telephone)	Audio recordings	Coding: • Bespoke system (coding for consultation characteristics postulated to influence patient recall) Other: • Measured duration and whether more than one problem was discussed
Stredler-Brown (2017)	Early intervention for children who are deaf or hard of hearing	USA	Video	16 clients and their parents Comparisons were made with existing findings from the literature	Video recordings	Coding • The presence of four behaviors (observation, direct instruction, parent practice with feedback, child behavior with provider feedback) used to coach parents of children who are deaf or hard of hearing
Tachakra and Rajani (2002)	Minor accident and treatment service	UK	Video	60 patients (30 face-to-face, 30 video)	Video recordings	Coding: • Consultation statistics (mean number of turns, words, interruptions, backchannels, and repairs)
Qualitative studies						
Ekberg et al. (2019)	Speech and language therapy	Australia	Video	4 clients (2 face-to-face, 2 video) 11 participants total, but analysis focuses on 4	Video recordings	Conversation analysis, with a focus on the different ways that speech and language therapists use physical objects depending on communication modality
Hewitt et al. (2010)	Primary care	UK	Telephone	65 patients (33 face-to-face, 32 telephone)	Audio recordings	Conversation analysis, with a focus on, amongst other things, how multi-problem consultations are handled in the different communication modalities
Pappas and Seale (2009)	Televascular and telecardiology	UK	Video	10 patients (5 telecardiology, 5 televascular)Comparisons were made with data extracts from Heath (1981, 1986)	Video recordings	Conversation analysis, with a focus on the opening phase of the consultation
Shaw et al. (2020)	Diabetes, antenatal diabetes, cancer surgery, heart failure	UK	Video	Diabetes: • 18 patients (7 face-to-face, 12 teleconsultations)Antenatal diabetes: • 12 patients (6 face-to-face, 6 teleconsultations)Cancer: • 18 patients (6 face-to-face, 12 teleconsultations)Heart failure: • 16 patients (9 face-to-face, 7 teleconsultations)	Video recordings (teleconsultations), audio recordings (face-to-face consultations)	Linguistic ethnography and conversation analysis, with a focus on openings and closings, physical examinations, problems with technology, and turn-taking
Stommel et al. (2019)	Post-operative consultations	Netherlands	Video	39 patients (17 face-to-face, 22 video)	Video recordings	Conversation analysis, with a focus on how surgeons display other-attentiveness at the beginning of consultations
Stommel et al. (2020)	Post-operative consultations	Netherlands	Video	39 patients (17 face-to-face, 22 video)	Video recordings	Conversation analysis, with a focus on the different ways that wound assessments are carried out between the two modalities
Mixed methods						
Shaw et al. (2018)	Diabetes, antenatal diabetes and cancer surgery	UK	Video and audio	Adult/young adult diabetes: • 12 patients (6 face-to-face, 6 teleconsultations)Antenatal diabetes: • 12 patients (6 face-to-face, 6 teleconsultations)Cancer surgery: • 10 patients (5 face-to-face, 5 teleconsultations)	Video recordings (teleconsultations), audio recordings (face-to-face consultations)	Quantitative coding: • RIASQuantitative other: • Measured duration and verbal dominanceQualitative: Although the primary focus is on the coding, the authors do include transcript extracts and qualitative observations

Note that if comparisons were done as part of a wider study or project (e.g. if interviews or questionnaires were collected alongside the recorded communication data), only data relevant to the comparisons is included in this table.

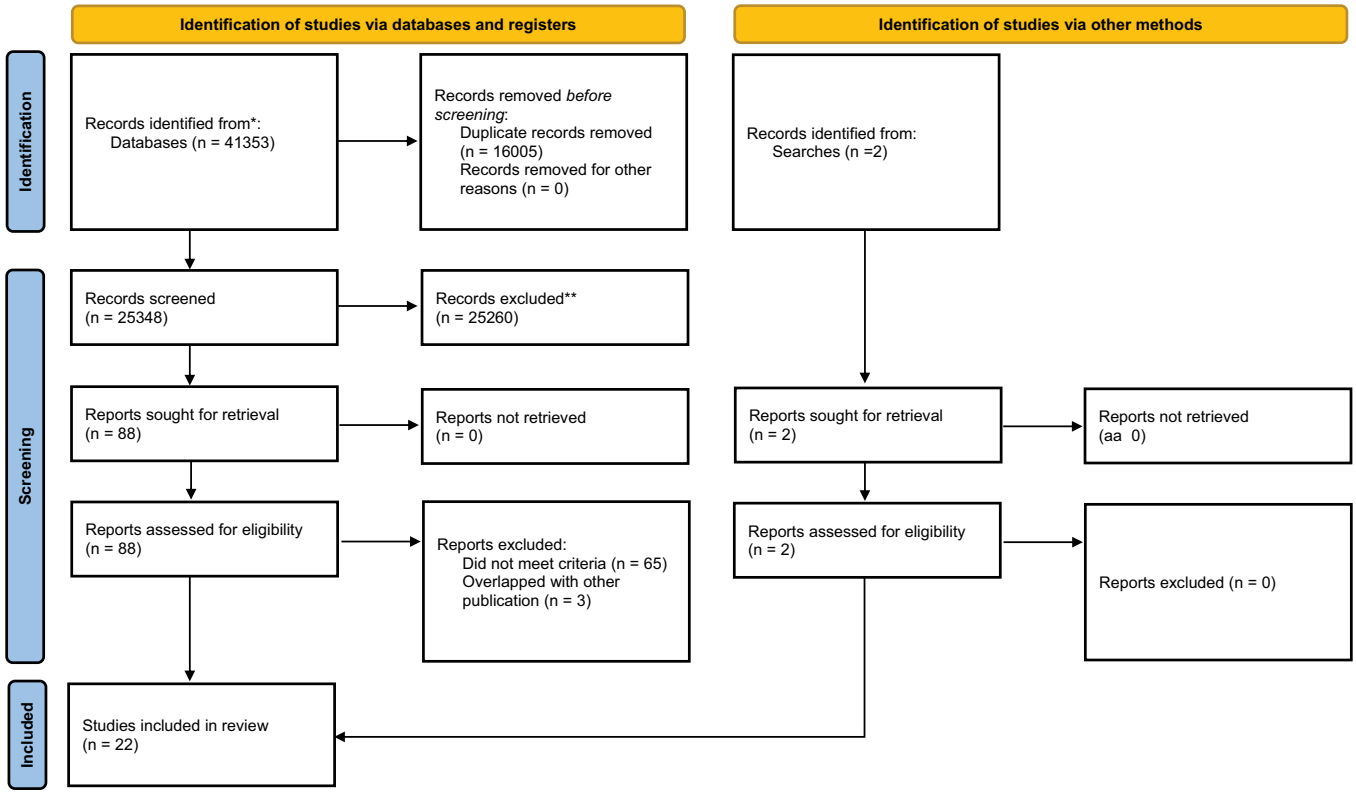

Figure 1. A flow diagram showing the paper screening process (from Page et al., 2021; visit: http://www.prisma-statement.org).

coding studies highlighting the prevalence and distribution of interactional behaviors and qualitative analyses providing insights into how these behaviors are manifested (see Robinson, 2007). There was also a lot of variety among the quantitative studies. For this reason, the findings are presented as a narrative synthesis, a suitable approach when "the review question dictates the inclusion of a wide range of research designs, producing … findings for which other approaches to synthesis are inappropriate" (Popay et al., 2006, p. 7).

We first created a list of all interactional variables analyzed in the coding studies. We then looked for patterns among the differences that had been identified. Finally, we considered how interactional behaviors observed in qualitative studies related to these patterns.

Results

Quality appraisal

Quality appraisal revealed much methodological variability within the studies reviewed. This was particularly true of quantitative studies, where there was a distinction between those studies that took a descriptive approach to comparing modalities (Innes et al., 2006; Laflamme et al., 2005; Stredler-Brown, 2017; Tachakra & Rajani, 2002) and those that used inferential statistics (Agha et al., 2010; Agha, Roter, et al., 2009; Ball et al., 1995; Day & Schneider, 2002; Demiris et al., 2005; Edison et al., 2013; Frueh et al., 2007; Hammersley et al., 2019; Liu et al., 2007; McKinstry et al., 2010, 2011). Virtually all studies used a preexisting, validated coding system, with the most common being the Roter Interaction Analysis System (RIAS) (Agha et al., 2010; Agha, Roter, et al., 2009; Hammersley et al., 2019; Innes et al., 2006; McKinstry et al., 2010; Shaw et al., 2018). Other systems included the Vanderbilt Psychotherapy Process Scale (Day & Schneider, 2002) and the Video Interactive Techniques for Automated Scoring System (Ball et al., 1995). Other studies used coding systems that were inspired by or adapted from larger systems (Demiris et al., 2005; Edison et al., 2013; Liu et al., 2007), while others used bespoke systems designed to address study-specific questions (Laflamme et al., 2005; McKinstry et al., 2011; Stredler-Brown, 2017).

Studies also varied in how "naturally occurring" the data were. While qualitative studies all made use of naturally occurring data, coding studies varied in their degree of manipulation: some studies recruited patients and randomly assigned them to face-to-face and telemedicine conditions; others allowed patients to choose a condition; and others, much like the qualitative studies, recorded naturally occurring consultations that had already been scheduled. Generally, this manipulation did not extend into the interactions themselves. Ball et al. (1995), however, did note that participants "were asked to perform tasks that would be part of the normal routine on the ward" (p. 23), suggesting some degree of involvement in how the interaction unfolded. There were also studies (e.g. Liu et al., 2007) where the same participants took part in both face-to-face and telemedicine conditions.

Most studies collected both the face-to-face and telemedicine data, with only a few comparing newly collected teleconsultations with existing face-to-face consultations. Although

there were cases where recordings/transcripts were only partially coded (e.g. Day & Schneider, 2002), the majority of quantitative studies coded full consultations/transcripts. It should also be noted that not all studies were focused on the comparative aspect; Innes et al. (2006), for example, compare their findings with those of an earlier study as a way to "stimulate reflection" rather than treating it as a core part of their study.

CASP discourages the use of scoring when using its appraisal tools (see, for example, CASP, 2018a), so we have not done so in this case. Nor have we excluded studies based on the appraisal alone. However, when synthesizing the evidence, we have used the appraisal process to determine which studies should be given greater prominence. With quantitative studies, this means that we have prioritized studies that used preexisting and validated coding systems (either in full or adapted form), that compared data newly collected data in both modalities, and that supported any comparisons with inferential statistics. Where necessary, we have noted these distinctions within the text.

Findings

Quantitative comparisons, which made up the majority of the studies, overwhelmingly suggested that face-to-face and teleconsultations were comparable. This can be illustrated, albeit in a crude way, by focusing on comparisons which tested for significance. Out of 189 comparisons across ten studies, 143 (76%) revealed no significant difference at all. (Note that this does not include every quantitative study – see Appendix for details.)

As overwhelming as this finding is, there were still four points of note among the comparisons of modalities. Firstly, there was consistent evidence that face-to-face consultations were on average longer than teleconsultations. It was also observed that the proportion of differences that were identified were more prevalent in primary than secondary care settings and were more likely to favor face-to-face over teleconsultations. Finally, even in the absence of substantial quantitative differences between modalities, there were still important qualitative differences. We will now discuss each of these points in turn.

Duration

The most straightforward and common way of measuring duration was the average consultation length in minutes or seconds. There were ten studies which used this metric (Agha et al., 2010; Agha, Roter, et al., 2009; Demiris et al., 2005; Edison et al., 2013; Hammersley et al., 2019; Laflamme et al., 2005; Liu et al., 2007; McKinstry et al., 2010, 2011; Shaw et al., 2018), with one of these studies (Shaw et al., 2018) comparing duration across three settings and one (Hammersley et al., 2019) comparing face-to-face to both video and telephone consultations. Of this total of 13 individual comparisons, 11 showed teleconsultations to be shorter on average (Agha et al., 2010; Demiris et al., 2005; Edison et al., 2013; Hammersley et al., 2019; Liu et al., 2007; McKinstry et al., 2010, 2011; Shaw et al., 2018) and two showed teleconsultations to be longer (Agha, Roter, et al., 2009; Laflamme et al., 2005).

Of the 10 comparisons which also tested for statistical significance, five showed face-to-face consultations to be significantly longer (Agha et al., 2010; Demiris et al., 2005; Liu et al., 2007; McKinstry et al., 2010, 2011) and five showed that the difference was not significant (Agha, Roter, et al., 2009; Edison et al., 2013; Shaw et al., 2018). There were no statistically significant comparisons showing teleconsultations to be longer (although Tachakra and Rajani (2002), for example, provide descriptive evidence of this being the case in an emergency treatment setting).

These findings on duration are supported by four studies (all in British general practice) which found that patients tend to raise fewer problems in teleconsultations (Hammersley et al., 2019; Hewitt et al., 2010; McKinstry et al., 2010, 2011). Hammersley et al. (2019), McKinstry et al. (2010), and McKinstry et al. (2011) showed these differences to be statistically significant, though only for video consultations in the case of Hammersley et al. Hewitt et al. did not test for significance, though did provide qualitative insights into this issue (see below).

Another way of measuring how time was distributed within consultations was the so-called dominance ratio, calculated by dividing the doctor's speech by the patient/companion's speech. Out of six such comparisons across three studies (Agha et al., 2010; Agha, Roter, et al., 2009; McKinstry et al., 2010; Shaw et al., 2018), there were two notable findings. Shaw et al. (2018) found that clinicians were significantly less dominant in antenatal teleconsultations, which the authors attribute to patients being asked to read out their blood glucose readings (in face-to-face consultations, the practitioner could see the readings directly). In a noninferiority study of three secondary care settings (pulmonary, endocrine, and rheumatology), Agha et al. (2010) found that physicians spoke more in face-to-face consultations. They could not confirm non-inferiority for this variable.

Although Edison et al. (2013) did not calculate a dominance ratio, they did note that patients and companions spoke significantly fewer words in remote than in face-to-face dermatology consultations.

Primary vs. secondary care

Although comparisons across the board suggested that (apart from duration) there were few significant differences between modalities, the extent to which this was true differed between studies in primary and secondary care. This can be highlighted again using our table of comparisons (see Appendix), where, excluding duration and related variables, only nine (20%) out of 46 significant differences identified were in secondary care. (Again, this table does, for practical reasons, exclude a large study of secondary care (Shaw et al., 2018). As we shall see below, however, this study identified few differences in the settings studied, making it broadly in line with this pattern.)

In some cases, studies in secondary care settings identified no differences at all. This was true of Ball et al. (1995), who, using the VRM (Stiles Verbal Response Modes) and VITAS (Video Interactive Techniques for Automated Scoring System)

coding schemes on psychiatric consultations, observed no significant differences in either verbal behaviors (interpretations, reflections, advisements, and questions) or non-verbal behaviors.

When differences were identified in secondary care settings, they tended to be few and inconsistent across studies. Agha, Roter, et al. (2009), in their study of pulmonology interactions in a veterans' hospital, found only one statistically significant difference (the number of orientation statements used by physicians) for either physicians or patients out of 10 communication variables coded using the Roter Interaction Analysis System. In dermatology, Edison et al. (2013), using a hybrid coding system based on the Davis Observational Codes, found one statistically significant difference: in the degree to which clinicians offered positive reinforcement (this was a subvariable of "Promoting compliance," which did not as a whole differ between modalities). Also in dermatology, Demiris et al. (2005) used this same system and identified only one variable ("Small talk") which showed any significant difference between modalities. Frueh et al. (2007) compared counseling using a protocol designed to evaluate therapists' adherence to and competence with a CBT treatment. They found no differences in the extent to which therapists adhered to this treatment, and only one difference in their (coderrated) competence (the skill with which they "introduced and explained new homework items," p. 863).

The single largest study of secondary care was Shaw et al. (2018), who compared RIAS communication variables in three different settings: adult/young adult diabetes, antenatal diabetes, and cancer surgery. They identified a small number of statistically significant differences across these settings, some of which were setting specific (differences in personal talk, clinician direction-giving, the number of requests, and the ratio of closed to open questions were all present in antenatal diabetes only), some of which applied across all three settings (patients' verbal attentiveness). The authors themselves note, however, that such differences "were rarely attributable to video-mediated interaction per se, but rather to contextual differences *beyond* the video link" (Shaw et al., 2018, p. 85). As they conclude, apart from "technology-related talk" (see below), "the kinds of talk were broadly similar" in face-to-face and telemedicine (p. 86).

This is not to say that there were no noteworthy differences in secondary care. In psychotherapy, for example, Day and Schneider (2002) observed differences in client participation for both telephone and video consultations, while Agha et al. (2010) found differences in psychosocial information exchange, emotional responsiveness, and partnership building. Overall, however, there was little evidence of statistically significant differences between modalities in secondary care, either in general or in specific settings.

While this was broadly true of primary care settings as well, the picture was more complex. There were four studies of primary care which tested for statistical significance: three in British general practice (Hammersley et al., 2019; McKinstry et al., 2010, 2011) and one in Japanese internal medicine (Liu et al., 2007). In British general practice, McKinstry et al. (2010) and Hammersley et al. (2019), who both used the RIAS, found similar differences in doctors' data gathering, patients' information giving, and doctors' patient education and counseling. These differences held whether face-to-face was compared to telephone alone (McKinstry et al.) or telephone and video (Hammersley et al.). With the exception of duration (see above), this kind of strong, consistent difference across studies was not found in secondary care settings, even when studies focused on the same specialty (e.g. Demiris et al. (2005) and Edison et al. (2013)).

Both McKinstry et al. (2010) and Hammersley et al. (2019) also identified differences in rapport building, but these differences were inconsistent (see below). Furthermore, they found, individually, differences in variables such as doctors' partnership building in video consultations (Hammersley et al.), as well as sub-categories of various overarching variables. Both studies also compared consultations based on Royal College of General Practitioner (RCGP) quality indicators/criteria, with Hammersley et al. identifying three differences and McKinstry et al. only identifying one. (Note, also, that the indicators used differed in some respects between the two studies.)

In Japanese internal medicine, Liu et al. (2007), using a different coding system, identified significant differences in empathy, praise, and facilitation utterances. McKinstry et al. (2011), meanwhile, found only one significant difference (the repetition of advice) between telephone and face-to-face consultations in general practice. It should be noted, however, that this study was not focused on communication per se but on "consultation features postulated to influence [patients'] recall," (p meaning that they did not use the RIAS and coded a smaller, more targeted range of behaviors.

While studies in primary care were broadly in line with our main finding therefore (i.e. with the exception of Liu et al. (2007), they showed more non-significant findings than significant ones), taken together, the majority of significant differences were clustered in this setting.

Direction of difference

In addition to considering the areas of medicine in which differences were identified, it is also important to consider the direction of the findings – that is, did differences reflect positively on face-to-face or telemedicine? Here again there was a clear pattern. Excluding duration and related variables, out of 46 statistically significant differences shown in the Appendix, only seven (15%) suggest an advantage in telemedicine.

As was already suggested in the previous section, there was little cross-study consistency in these differences. This was particularly true of those differences which showed an advantage in telemedicine. Demiris et al. (2005), for example, found evidence of greater levels of small talk in teleconsultations, but this finding was not supported by Edison et al. (2013) comparable study in the same setting, which showed no significant difference in informal talk. In primary care, Hammersley et al. (2019) found that rapport building was significantly higher in both telephone and video consultations; other primary care studies suggested the opposite, however, with significantly *lower* levels of rapport building (McKinstry et al., 2010) and empathy (Liu et al., 2007) in teleconsultations. Day and Schneider (2002), using the

Vanderbilt Psychotherapy Process Scale, observed greater levels of client participation in both audio-only and video teleconsultations, a finding which they acknowledge is surprising and which they attribute to patients potentially wanting to "get their voices heard when technology came between them and their therapists" (p. 502); however, there was little evidence of such an advantage elsewhere. And while Shaw et al. (2018) did, in their study of three secondary care settings, observe several significant differences in favor of telemedicine, as was noted above, the authors themselves generally downplay the import of these findings. For example, in observing that there was significantly more personal talk in virtual antenatal diabetes interactions, they note that this finding was "probably spurious" and perhaps explained by personal talk in face-to-face encounters occurring off-camera "in the waiting area and during the walk to the waiting room" (Shaw et al., 2018, p. 74).

There was some cross-study consistency in findings which showed face-to-face consultations to be superior. This was true of the findings in British general practice relating to clinicians' data gathering, patients' information giving, and clinicians' patient education and counseling (Hammersley et al., 2019; McKinstry et al., 2010), which consistently favored face-to-face consultations across patients and doctors and, in the case of Hammersley et al., both telephone and video remote consultations. Even these findings were not supported by Liu et al. (2007), however, who, applying a different coding system in Japanese internal medicine, found no significant difference in clinicians' questioning. These findings are partly supported in secondary care by Agha et al. (2010), who, as part of a noninferiority study, found lower levels of psychosocial (though not biomedical) information gathering in pulmonary, endocrine, and rheumatology clinics at a veterans' hospital. In nursing home encounters, meanwhile, Laflamme et al. (2005) show that residents indicated understanding of their treatment plan in substantially fewer video consultations (38%) than face-to-face to consultations (100%), although there was no inferential statistical comparison.

Yet regardless of patterns within individual communication variables or settings, there was a clear tendency *in the aggregate* for those differences that were identified to favor face-to-face consultations over teleconsultations. This is of course in line with the findings on duration discussed above.

Qualitative differences

To summarize so far, we have shown that, outside of a small number of domains, quantitative comparative studies overwhelmingly suggested that there is little significant difference between face-to-face and telemedicine. To fully understand how these modalities compare, however, we also need to consider qualitative comparative studies, which can highlight aspects which are inaccessible to a quantitative approach.

As alluded to above, Hewitt et al. (2010) provide a good example of this. Where other studies provided statistical evidence that primary care consultations tend to be more mono-topical (Hammersley et al., 2019; McKinstry et al., 2010), Hewitt et al. also highlight the interactional dynamics through which this occurs. They show how both patients and doctors can treat consultations as mono-topical from the start, and how doctors, following verbal examination, can invite the patient to a face-to-face consultation (where, presumably, other health matters could be raised). They also note how telephone consultations had fewer gaps in which patients could raise additional health matters (see also Shaw et al., 2018) or engage in small talk.

Qualitative research could also explore the impact of technology and technology related talk. While two quantitative studies did code for "technology talk" (Demiris et al., 2005) and "technical issues" (Edison et al., 2013), one of these studies only coded for it in teleconsultations because it was "obviously not applicable to actual visits" (Demiris et al., 2005, p. 854) and the other coded for it in both modalities while acknowledging that it "was not statistically comparable because the technical issues during the teledermatology visits entailed dealings with camera or lighting" (Edison et al., 2013, p. 511). Some coding studies did provide an indirect insight into the problems potentially caused by technology by highlighting significantly larger numbers of requests for repetition or unintelligible utterances in remote consultations (Agha, Roter, et al., 2009; Liu et al., 2007). Shaw et al. (2018), also, provide a detailed quantitative breakdown of different types of technology talk (e.g. socioemotional), which they supplement with qualitative observations (see below). Overall, however, comparative coding approaches clearly struggled with a variable that was intrinsically more relevant to one modality than the other.

This was not the case with qualitative studies, which could provide microanalytic insights into how technology altered the interaction. As noted above, for example, Shaw et al. (2018) show how there is a period of adjustment at the beginning of remote consultations as a connection is established. It is only after this point "that the doctor asks *how are you*, signaling the start of the clinical consultation and enabling the social rituals of *being a doctor* and *being a patient* to begin." This finding is supported by Stommel et al. (2019), who also show how the question "how are you?" is routinely used by doctors to transition between testing technology and the main business of the consultation in post-surgery consultations. It is also supported by Pappas and Seale (2009), who demonstrate how, unlike the recurring structure exhibited by the openings of face-to-face consultations, telecardiology and televascular consultations do not begin in any systematic way, with participants working to attune themselves to the new mode of communication. Shaw et al. (2018), furthermore, make a similar point about the end of consultations, noting that while remote consultations tended to conclude in a definitive way, face-to-face consultations tended to have less well-defined endings (e.g. because patients were asked to wait to speak to someone else).

Qualitative studies can also provide insight into how communication behaviors manifest in each modality. This is in contrast to coding studies which focus on whether and to what extent a given behavior is present. Stommel et al. (2020), for example, analyzed face-to-face and video-mediated postoperative consultations, with a specific focus on how closed wound assessments were conducted. They showed that face-to-face wound assessments were, with one exception, conducted visually, with the doctor making firm assessments based on the visual evidence

(e.g. "looks neat"). In contrast, telemedicine assessments were overwhelmingly carried out verbally (even though video calls made a visual assessment theoretically possible), with patients taking a primary role in assessing their own wounds and doctors arriving at "qualified wound assessments . . . grounded in the patient's evaluation rather than their own observation or examination." While both these behaviors could be coded as "wound assessment" (or "examination" etc.), the form differed between modalities. Related to this, Shaw et al. (2020) show how physical examinations that presented no problem in a face-to-face consultation could be challenging remotely, where "success was often dependent on the type of technology . . . , the presence of a third party who could assist the patient, the patient's mobility, and the technological competence of the participants."

Another example of this kind of analysis can be seen in Ekberg et al. (2019), who compared how speech and language therapists (SLTs) used toys and other physical objects in face-to-face and teleconsultations with children. They suggest that, in face-to-face consultations, SLTs used physical objects as a way of incentivizing children to produce a target utterance, restricting access to the object until the child did so. In telemedicine consultations, of course, this was not an option; instead, the SLTs themselves used physical objects to facilitate children's engagement in the therapy. Again, the same behavior was adapted according to the modality.

Taken together, therefore, this qualitative evidence shows that as well as differing in how much they manifest between modalities, communication behaviors can also differ in how they manifest.

Discussion

The aim of this article was to review the comparative literature to identify which differences, if any, exist between face-to-face and teleconsultations.

Quantitative coding studies overwhelmingly suggested that there were few significant differences between modalities. The one exception to this was duration, with studies across multiple settings suggesting that face-to-face consultations were longer on average than their remote counterparts (although this difference did not always meet the threshold of statistical significance). This is in line with an earlier review of comparative studies of therapy interaction, which concluded that, apart from duration, there was "a lack of support for arguments that the telephone has a detrimental effect on interactional aspects of psychological therapy" (Irvine et al., 2020, p. 129).

Even in the absence of other such cross-setting differences, there were some patterns of note. We observed that differences were more commonly identified in primary care, with findings in secondary care settings tending to be few and inconsistent. This pattern could be attributed to secondary care settings typically having a tighter focus and more predictable structure, making them easier to adapt to telemedicine.

We also observed that those differences that were identified tended to favor face-to-face consultations over teleconsultations. Indeed, there was some cross-study evidence that there was significantly more information exchange and patient education and counseling in face-to-face consultations, though this finding was limited to British primary care.

Our findings can be interpreted in two ways. Proponents of teleconsultations could point to the overwhelming lack of significant differences between modalities as support for their continued use. Certainly, this finding is in line with evidence from interview and survey research showing generally high satisfaction (Miller, 2001; Nguyen et al., 2020). Critics, however, could focus on the qualitative differences and the finding that, in the aggregate, those quantitative differences that were identified tended to favor face-to-face consultations. Advocates could argue in turn that the studies we reviewed were predisposed to such a finding because they were often using tools that were originally developed for face-to-face consultations and were thus measuring teleconsultations by a nonobjective standard that they could never possibly meet (see Miller and Nelson (2005) and Nelson et al., (2010) on modifying the Roter Interaction Analysis System for application to teleconsultations). The implication of this possibility, that teleconsultations have interactional benefits to which current approaches are not attuned, is an intriguing one and a possible direction for future research.

Directions for future research

There were three notably understudied areas that could form the basis of future research.

1. Home-based telemedicine

Of the 22 studies reviewed, eight involved participants traveling somewhere (e.g. a local hospital) to participate in teleconsultations; that is, they would be in one room while the practitioner would be in another. In some cases, this was done so that the research team could control extraneous variables (e.g. Frueh et al., 2007); in other cases, this was simply how the telemedicine system in question operated at the time of the research (e.g. Demiris et al., 2005). This was especially notable in the coding-only studies, with only one of the qualitative studies involving patients traveling to a different location to participate in the telemedicine encounter. It is also worth noting that only one of these studies was conducted in the last ten years (Edison et al., 2013), with more recent studies overwhelmingly focusing on home-based telemedicine.

These older studies are not representative of telemedicine in the time of Covid, nor are they likely to be representative of telemedicine going forward. As Day and Schneider (2002) acknowledge, such research designs mean that one of the key benefits of telemedicine (that it allows patients to communicate flexibly, without needing to travel) is neglected. If a patient's location during a consultation does indeed impact on their communication (and the assumption of such an impact does of course underlie those studies where patients were asked to travel to a controlled setting to avoid extraneous variables that might arise at home), then much of the earlier comparative research on telemedicine may have a limited applicability to the present day.

2. Non-verbal behavior

A key concern about telemedicine is the impact on bodily communication and physical examination (Bulik, 2008; Miller, 2003). Indeed, in a 2001 review, Miller found that findings from observational telemedicine studies were generally positive except when it came to "non-verbal behavior and lack of touch" (p. 12). Despite this, these topics were largely absent from the comparative studies reviewed here.

This was particularly true of coding studies, for several reasons. Firstly, certain coding studies used audio recordings even when video consultations were being analyzed. When "nonverbal communication" was analyzed, furthermore, this could mean tone of voice rather than embodied conduct (e.g. Agha, Roter, et al., 2009). Indeed, the most used coding system was the RIAS, which, as Miller and Nelson (2005) note, is primarily focused on verbal content, taking into account "laughter, crying, voice intonation, and tone" but *not* "nonverbal behaviors such as eye contact, gaze, posture, facial expressions, body positioning, proximity, touch, activity … and other cues that modify the meaning of verbal utterances" (p. 51) (see Sandvik et al. (2002) for a more general critique of the RIAS). Finally, even if coding studies did consider bodily communication, it was subsumed into a broader category ("information gathering," "empathy" etc.). The only exception to this was Ball et al. (1995), who coded and presented separate results for "angle of recline" and "mutual gaze."

Some qualitative studies did offer detailed analyses of how, for example, physical examinations are conducted remotely (Shaw et al., 2020). This was not the case across the board, however, with another qualitative study noting that analyzing "visual cues" was "not feasible because the analysts did not always have a view of the proceedings of all participants" (Pappas & Seale, 2009, p. 1231).

Overall, the comparative impact of telemedicine on bodily communication remains underexplored in the existing observational literature. Again, the lack of focus on topic may be a product of shifts in culture and technology; nonverbal behavior might not have been relevant when the telephone was the dominant form of telemedicine, but the widespread availability and adoption of videoconferencing brought about by COVID-19 is likely to change this.

3. Third parties

Third parties (i.e. friends, family members, and others who participate in consultations outside of the core patient-professional dynamic) can play a crucial role in medical interaction (e.g. Clayman et al., 2005). This is particularly true when, for example, the patient is suffering from cognitive impairment or when the reason for the consultation is a loss consciousness, meaning that the third party can provide a more reliable account than the patient themselves. This role has rarely been discussed in the comparative literature. Of the 22 studies reviewed, nine did not mention third parties at all. The remaining 13 studies did allude to some third-party presence, whether that was an additional staff member or family and friends. However, most of these allusions were in passing, with no analytic focus specifically on the third parties.

Coding studies would often group third party communication with either the doctor or the patient (e.g. Edison et al., (2013), where physician and resident communication was grouped together as "physician communication" and the talk of everyone else in the consultation was grouped together as "other words"). This issue is noted by Miller and Nelson (2005), who suggest that the RIAS (again, the most used coding system in the studies reviewed) be modified to take a more nuanced approach to third-party contributions in telemedicine interactions. These issues were not limited to coding studies, however, with qualitative studies sometimes presenting transcript extracts featuring only doctor-patient communication, with no contributions from others. In short, there was an overwhelming tendency to present interaction as dyadic even when others were present.

The lack of focus on third parties is understandable, inasmuch as research in this area is still exploratory and novel. Going forward, however, it will be important to focus on third parties because, as existing research has shown, third parties are not simply extensions of the patient or provider – they have their own distinct roles to play. This is highlighted by those rare comparative studies that did focus on third parties. Tachakra and Rajani (2002), for example, compared doctor-patient and nurse-patient communication separately, and found some differences between the two (although they did not statistically compare these differences). Shaw et al. (2020), meanwhile, highlight the role that family members can play in assisting with examination when the patient and provider are do not share the same physical space.

Future research should thus compare third party communication separately from the patient-provider dyad to determine how it is impeded or facilitated when communicating remotely. Such research could also compare how providers manage the participation of friends and family members in teleconsultations, given that this is often done non-verbally (e.g. using gaze) in face-to-face consultations (Tiitinen & Ruusuvuori, 2012).

4. Communication outcomes

Although the studies reviewed here often measured patient satisfaction alongside the analysis of interactional data, studies which measured other clinical outcomes (e.g. engagement, symptom resolution) were rare (for an exception, see Day and Schneider (2002), who measured symptom severity, global functioning, and other clinical outcomes). Given that earlier research in face-to-face settings has shown that clinical communication can influence outcomes (see Street et al. (2009) on the seven "pathways" between communication and health outcomes), it would make sense for future comparative studies to measure whether observed interactional differences between modalities also affect outcomes between these modalities.

Limitations of this review

Our criteria for this review were inclusive, taking into account any study featuring a communicative comparison between face-to-face and teleconsultations in a medical setting. This has allowed us to present a wider overview of comparative

studies, highlighting the topics and findings that cut across this subsection of research.

However, the heterogeneity of the studies does make it difficult to determine whether any differences identified are related to the communication modality or to other variables. Such variables include the medical specialty in which the communication occurred, the type of telemedicine compared, the country in which the interactions took place, whether there was an existing relationship between patient and practitioner, participants' degree of familiarity and expertise with telemedicine technology, whether consultations studied were first-time or follow-ups, and, in the case of coding studies, the system used.

Such variability would likely be an issue for this type of review even if our criteria had been more targeted; this is an inevitable part of reviewing studies from an area of research that remains new and exploratory. As Miller (2010) notes, "[i]t is difficult to generalize across the ... interaction analysis studies conducted to date" (p. 2) because of the diversity of these studies. Nonetheless, our findings are notably varied, and we cannot rule out the possibility that focusing on a particular subset of studies within those retrieved might have identified more systematic patterns. We should also emphasize that just because certain findings are described in some papers but others, that does not mean that the paper which described them was wrong or that the findings are not relevant.

All the studies reviewed here predate the COVID-19 pandemic in terms of data collection. This is important because pre-pandemic the use of telemedicine was limited to specific contexts and mostly to clinicians (and often also patients) who had a special interest in using it. This means, for example, that Hewitt et al. (2010) finding that general practitioners (GPs) tended to invite new telephone patients for a face-to-face consultation rather than prescribing right away could be different now that a face-to-face consultation would involve potential exposure to the virus. Similarly, Pappas and Seale's (2009) observations about participants attuning themselves to technology that was relatively new at the time may no longer apply. While our findings might be broadly representative of telemedicine communication when it is used by choice, then, it is unclear how reflective they would be when face-to-face consultations are the exception rather than the norm.

Conclusion

Although there is some variability between specialties, face-to-face and teleconsultation communication appears to be comparable in most aspects studied so far. Although this finding supports the high rates of physician and patient satisfaction seen in other telemedicine research, the picture is not wholly positive. There are also important and potentially consequential aspects of medical interaction (third parties and bodily communication) that have not been widely studied in the existing comparative research, as well as the possibility that earlier research in this area does not reflect the reality of telemedicine in a post-Covid world. Qualitative studies, meanwhile, show that even if a behavior is technically present in both modalities, the form that it takes may differ. This suggests that comparative research on telemedicine communication is, much like the field of telemedicine itself, still in its early stages, with much work to be in broadening its scope.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was jointly funded by Epilepsy Research UK and Epilepsy Action through an Innovations in Healthcare Urgent Research grant.

ORCID

Joseph Ford http://orcid.org/0000-0002-2248-8591
Markus Reuber http://orcid.org/0000-0002-4104-6705

References

Agha, Z., McNutt, G., Roter, D., Laud, P., Schapira, R., & Gray, B. (2010, April 28 – May 1). Content analyses of physician-patient communication during telemedicine [Conference presentation abstract]. *Society of General Internal Medicine: 33rd Annual Meeting*, Minneapolis, MN, United States. https://doi.org/10.1007/s11606-010-1338-5

Agha, Z., Roter, D. L., & Schapira, R. M. (2009). An evaluation of patient-physician communication style during telemedicine consultations. *Journal of Medical Internet Research*, *11*(3), e36. https://doi.org/10.2196/jmir.1193

Agha, Z., Schapira, R. M., Laud, P. W., McNutt, G., & Roter, D. L. (2009). Patient satisfaction with physician-patient communication during telemedicine. *Telemedicine and E-Health*, *15*(9), 830–839. https://doi.org/10.1089/tmj.2009.0030

Bailey, S. C., O'Conor, R., Bojarski, E. A., Mullen, R., Patzer, R. E., Vicencio, D., Jacobson, K. L., Parker, R. M., & Wolf, M. S. (2015). Literacy disparities in patient access and health-related use of Internet and mobile technologies. *Health Expectations*, *18*(6), 3079–3087. https://doi.org/10.1111/hex.12294

Ball, C. J., McLaren, P. M., Summerfield, A. B., Lipsedge, M. S., & Watson, J. P. (1995). A comparison of communication modes in adult psychiatry. *Journal of Telemedicine and Telecare*, *1*(1), 22–26. https://doi.org/10.1177/1357633X9500100105

Beach, W. A. (2012). *Handbook of patient-provider interactions: Raising and responding to concerns about life, illness, and disease*. Hampton Press.

Brink-Muinen, A. V. D., van Dulmen, S., Messerli-Rohrbach, V., & Bensing, J. (2002). Do gender-dyads have different communication patterns? A comparative study in Western-European general practices. *Patient Education and Counseling*, *48*(3), 253–264. https://doi.org/10.1016/s0738-3991(02)00178-7

Bulik, R. J. (2008). Human factors in primary care telemedicine encounters. *Journal of Telemedicine and Telecare*, *14*(4), 169–172. https://doi.org/10.1258/jtt.2007.007041

Clayman, M. L., Roter, D., Wissow, L. S., & Bandeen-Roche, K. (2005). Autonomy-related behaviors of patient companions and their effect on decision-making activity in geriatric primary care visits. *Social Science & Medicine*, *60*(7), 1583–1591. https://doi.org/10.1016/j.socscimed.2004.08.004

Courtney, E., Blackburn, D., & Reuber, M. (2021). Neurologists' perceptions of utilising tele-neurology to practice remotely during the COVID-19 pandemic. *Patient Education and Counseling*, *104*(3), 452–459. https://doi.org/10.1016/j.pec.2020.12.027

Critical Appraisal Skills Programme. (2018a). *CASP case control study checklist*. https://casp-uk.net/casp-tools-checklists//

Critical Appraisal Skills Programme. (2018b). *CASP qualitative studies checklist*. https://casp-uk.net/casp-tools-checklists/

Critical Appraisal Skills Programme. (2020). *CASP randomised controlled trials checklist*. https://casp-uk.net/casp-tools-checklists

Day, S. X., & Schneider, P. L. (2002). Psychotherapy using distance technology: A comparison of face-to-face, video, and audio treatment. *Journal of Counseling Psychology*, *49*(4), 499–503. https://doi.org/10.1037/0022-0167.49.4.499

Demiris, G., Edison, K., & Vijaykumar, S. (2005). A comparison of communication models of traditional and video-mediated health care delivery. *International Journal of Medical Informatics*, *74*(10), 851–856. https://doi.org/10.1016/j.ijmedinf.2005.03.018

Edison, K. E., Fleming, D. A., Nieman, E. L., Stine, K., Chance, L., & Demiris, G. (2013). Content and style comparison of physician communication in teledermatology and in-person visits. *Telemedicine and E-Health*, *19*(7), 509–514. https://doi.org/10.1089/tmj.2012.0227

Ekberg, S., Danby, S., Theobald, M., Fisher, B., & Wyeth, P. (2019). Using physical objects with young children in 'face-to-face' and telehealth speech and language therapy. *Disability and Rehabilitation*, *41*(14), 1664–1675. https://doi.org/10.1080/09638288.2018.1448464

Frueh, B. C., Monnier, J., Grubaugh, A. L., Elhai, J. D., Yim, E., & Knapp, R. (2007). Therapist adherence and competence with manualized cognitive-behavioral therapy for PTSD delivered via videoconferencing technology. *Behavior Modification*, *31*(6), 856–866. https://doi.org/10.1177/0145445507302125

Gomez, T., Anaya, Y. B., Shih, K. J., & Tarn, D. M. (2021). A qualitative study of primary care physicians' experiences with telemedicine during Covid-19. *Journal of the American Board of Family Medicine*, *34* (Supplement), S61–70. https://doi.org/10.3122/jabfm.2021.S1.200517

Hammersley, V., Donaghy, E., Parker, R., McNeilly, H., Atherton, H., Bikker, A., Campbell, J., & McKinstry, B. (2019). Comparing the content and quality of video, telephone, and face-to-face consultations: A non-randomised, quasi-experimental, exploratory study in UK primary care. *The British Journal of General*, *69*(686), e595–604. https://doi.org/10.3399/bjgp19X704573

Harrison, R., Macfarlane, A., Murray, E., & Wallace, P. (2006). Patients' perceptions of joint teleconsultations: A qualitative evaluation. *Health Expectations*, *9*(1), 81–90. https://doi.org/10.1111/j.1369-7625.2006.00368.x

Heath, C. (1981). The opening sequence in doctor-patient interaction. In P. Atkinson & C. Heath (Eds.), *Medical work: Realities and routines* (pp. 71–90). Gower.

Heath, C. (1986). *Body movement and speech in medical interaction*. Cambridge University Press.

Heritage, J., & Maynard, D. W. (Eds.). (2006). *Communication in medical care: Interaction between primary care physicians and patients*. Cambridge University Press.

Heritage, J., Robinson, J. D., Elliott, M. N., Beckett, M., & Wilkes, M. (2007). Reducing patients' unmet concerns in primary care: The difference one word can make. *Journal of General Internal Medicine*, *22* (10), 1429–1433. https://doi.org/10.1007/s11606-007-0279-0

Hewitt, H., Gafaranga, J., & McKinstry, B. (2010). Comparison of face-to-face and telephone consultations in primary care: Qualitative analysis. *British Journal of General Practice*, *60*(574), e201–212. https://doi.org/10.3399/bjgp10X501831

Innes, M., Skelton, J., & Greenfield, S. (2006). A profile of communication in primary care physician telephone consultations: Application of the Roter Interaction Analysis System. *British Journal of General Practice*, *56*(526), 363–368. https://bjgp.org/content/73/729/148

Irvine, A., Drew, P., Bower, P., Brooks, H., Gellalty, J., Armitage, C. J., Barkham, M., McMillan, D., & Bee, P. (2020). Are there interactional differences between telephone and face-to-face psychological therapy? A systematic review of comparative studies. *Journal of Affective Disorders*, *265*, 120–131. https://doi.org/10.1016/j.jad.2020.01.057

Laflamme, M. R., Wilcox, D. C., Sullivan, J., Schadow, G., Lindbergh, D., Warvel, J., Buchanan, H., Ising, T., Abernathy, G., Perkins, S. M., Daggy, J., Frankel, R. M., Dexter, P., McDonald, C. J., & Weiner, M. (2005). A pilot study of usefulness of clinician-patient videoconferencing for making routine medical decisions in the nursing home. *Journal of the American Geriatric Society*, *53*(8), 1380–1385. https://doi.org/10.1111/j.1532-5415.2005.53422.x

Liu, X., Sawada, Y., Takizawa, T., Sato, H., Sato, M., Sakamoto, H., Utsugi, T., Sato, K., Sumino, H., Okamura, S., & Sakamaki, T. (2007). Doctor-patient communication: A comparison between telemedicine consultation and face-to-face consultation. *Internal Medicine*, *46*(5), 227–232. https://doi.org/10.2169/internalmedicine.46.1813

McKinstry, B., Hammersley, V., Burton, C., Pinnock, H., Elton, R., Dowell, J., Sawdon, N., Heaney, D., Elwyn, G., & Sheikh, A. (2010). The quality, safety and content of telephone and face-to-face consultations: A comparative study. *Quality & Safety in Health Care*, *19*(4), 298–303. https://doi.org/10.1136/qshc.2008.027763

McKinstry, B., Watson, P., Elton, R. A., Pinnock, H., Kidd, G., Meyer, B., Logie, R., & Sheikh, A. (2011). Comparison of the accuracy of patients' recall of the content of telephone and face-to-face consultations: An exploratory study. *Postgraduate Medical Journal*, *87*(1028), 394–399. https://doi.org/10.1136/pgmj.2010.101287

Miller, E. A. (2001). Telemedicine and doctor-patient communication: An analytical survey of the literature. *Journal of Telemedicine and Telecare*, *7*(1), 1–17. https://doi.org/10.1258/1357633011936075

Miller, E. A. (2003). The technical and interpersonal aspects of telemedicine: Effects on doctor-patient communication. *Journal of Telemedicine and Telecare*, *9*(1), 1–7. https://doi.org/10.1258/135763303321159611

Miller, E. A. (2010). *Telemedicine and the patient-provider relationship: What we know so far*. Nuffield Council's Working Party. https://www.nuffieldbioethics.org/wp-content/uploads/Miller-E-2010-Evidence-review-Telemedicine-and-the-Provider-Patient-Relationship-what-we-know-so-far.pdf

Miller, E. A., & Nelson, E. L. (2005). Modifying the Roter Interaction Analysis System to study provider-patient communication in telemedicine: Promises, pitfalls, insights, and recommendations. *Telemedicine Journal and E-Health*, *11*(1), 44–55. https://doi.org/10.1089/tmj.2005.11.44

Nakornchai, T., Conci, E., Hensiek, A., & Brown, J. W. L. (2021). Clinician and patient experience of neurology telephone consultations during the COVID-19 pandemic. *Postgraduate Medical Journal*, *98*(1161), 533–538. https://doi.org/10.1136/postgradmedj-2021-141234

Nelson, E. L., Miller, E. A., & Larson, K. A. (2010). Reliability associated with the Roter Interaction Analysis System (RIAS) adapted for the telemedicine context. *Patient Education and Counseling*, *78*(1), 72–78. https://doi.org/10.1016/j.pec.2009.04.003

Nguyen, M., Waller, M., Pandya, A., & Portnoy, J. (2020). A review of patient and provider satisfaction with telemedicine. *Current Allergy and Asthma Reports*, *20*(11), 72. https://doi.org/10.1007/s11882-020-00969-7

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systemaic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Pappas, Y., & Seale, C. (2009). The opening phase of telemedicine consultations: An analysis of interaction. *Social Science & Medicine*, *68*(7), 1229–1237. https://doi.org/10.1016/j.socscimed.2009.01.011

Peräkylä, A. (1998). Authority and accountability: The delivery of diagnosis in primary health care. *Social Psychology Quarterly*, *61*(4), 301–320. https://doi.org/10.2307/2787032

Pires, C. M., & Cavaco, A. M. (2014). Communication between health professionals and patients: Review of studies using the RIAS (Roter Interaction Analysis System) method. *Revista da Associacao Medica Brasileira*, *60*(2), 156–172. https://doi.org/10.1590/1806-9282.60.02.014

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews*. ESRC Methods Programme. https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf

Reuber, M., Toerien, M., Shaw, R., & Duncan, R. (2015). Delivering patient choice in clinical practice: A conversation analytic study of communication practices used in neurology clinics to involve patients in decision-making. *Health Services and Delivery Research*, *3*(7), 1–170. https://doi.org/10.3310/hsdr03070

Robinson, J. D. (2007). The role of numbers and statistics within conversation analysis. *Communication Methods and Measures*, *1*(1), 66–75. https://doi.org/10.1080/19312450709336663

Roter, D., & Larson, S. (2002). The Roter interaction analysis system (RIAS): Utility and flexibility for analysis of medical interactions. *Patient Education and Counseling*, *46*(4), 243–251. https://doi.org/10.1016/s0738-3991(02)00012-5

Sandelowski, M., Voils, C. I., & Barroso, J. (2006). Defining and designing mixed research synthesis studies. *Research in the Schools*, *13*(1), 29.

Sandvik, M., Eide, H., Lind, M., Graugaard, P. K., Torper, J., & Finset, A. (2002). Analyzing medical dialogues: Strength and weakness of Roter's interaction analysis system (RIAS). *Patient Education and Counseling*, *46*(4), 235–241. https://doi.org/10.1016/s0738-3991(02)00014-9

Shaw, S. E., Seuren, L. M., Wherton, J., Cameron, D., A'Court, C., Vijayaraghavan, S., Morris, J., Bhattacharya, S., & Greenhalgh, T. (2020). Video consultations between patients and clinicians in diabetes, cancer, and heart failure services: Linguistic ethnographic study of video-mediated interaction. *Journal of Medical Internet Research*, *22*(5), e18378. https://doi.org/10.2196/18378

Shaw, S., Wherton, J., Vijayaraghavan, S., Morris, J., Bhattacharya, S., Hanson, P., Campbell-Richards, D., Ramoutar, S., Collard, A., Hodkinson, I., & Greenhalgh, T. (2018). *Advantages and limitations of virtual online consultations in a NHS acute trust: The VOCAL mixed-methods study*. NIHR Journals Library.

Stommel, W., van Goor, H., & Stommel, M. (2019). Other-Attentiveness in video consultation openings: A conversation analysis of video-mediated versus face-to-face consultations. *Journal of Computer-Mediated Communication*, *24*(6), 275–292. https://doi.org/10.1093/jcmc/zmz032

Stommel, W. J. P., van Goor, H., & Stommel, M. W. J. (2020). The impact of video-mediated communication on closed wound assessments in postoperative consultations: Conversation analytical study. *Journal of Medical Internet Research*, *22*(5), e17791. https://doi.org/10.2196/17791

Stredler-Brown, A. (2017). Examination of coaching behaviors used by providers when delivering early intervention via telehealth to families of children who are deaf or hard of hearing. *Perspectives of the ASHA Special Interest Groups*, *2*(9), 25–42. https://doi.org/10.1044/persp2.SIG9.25

Street, R. L., Jr., Makoul, G., Arora, N. K., & Epstein, R. M. (2009). How does communication heal? Pathways linking clinician-patient communication to health outcomes. *Patient Education and Counseling*, *74*(3), 295–301. https://doi.org/10.1016/j.pec.2008.11.015

Tachakra, S., & Rajani, R. (2002). Social presence in telemedicine. *Journal of Telemedicine and Telecare*, *8*(4), 226–230. https://doi.org/10.1258/135763302320272202

Tiitinen, S., & Ruusuvuori, J. (2012). Engaging parents through gaze: Speaker selection in three-party interactions in maternity clinics. *Patient Education and Counseling*, *89*(1), 38–43. https://doi.org/10.1016/j.pec.2012.04.009

Appendix

The following table summarizes the comparisons done across various studies included in the review. It highlights whether comparisons favored face-to-face, telemedicine, or showed no significant difference. It also highlights whether these differences were in variables related to the duration and distribution of talk (duration, dominance etc.) or variables related to the content of communication.

Only comparisons that tested for significance are included. Furthermore, two studies that did test for significance are excluded. In the case of Agha et al. (2010), this was because the study used a non-inferiority design; in the case of S. Shaw et al. (2018), it was because we could not determine the total number of comparisons that had been conducted.

Note that when a variable had sub-variables only the main variable has been counted. Please consult the original publications for full details. Some studies (e.g. McKinstry et al., 2011) compared other variables alongside communication variables. These other variables have not been included in the counts below.

Paper	Comparisons favoring FtF	Comparisons favoring TM	No significant difference	Total no. of comparisons
Primary care				
Liu et al. (2007) [Japanese internal medicine]	9 Duration and dominance[1]: 5 Other: 4 *Longer (according to several measures of duration, including time, utterances, and conversational turns), with more empathy, praise, and facilitation.*	0	5 Duration and dominance: 2 Other: 3 *No differences in certain measures of duration (e.g. utterances per minute), simultaneous utterances, and questioning (open and closed).*	14 Duration and dominance: 7 Other: 7
McKinstry et al. (2010) [British general practice] [Based on adjusted *p* values]	9 Duration and dominance: 2 Other: 7 *Longer, with more problems discussed, information exchange, patient education/counseling, direction giving, and rapport building.*	0	21 Duration and dominance: 1 Other: 20 *No differences in verbal dominance, patient-centeredness/involvement, and partnership building/disagreement. All but one RCGP quality criteria non-significant.*	30 Duration and dominance: 3 Other 27
McKinstry et al. (2011) [British general practice]	3 Duration and dominance: 2 Other: 1 *Longer, with more multi-problem consultations and repetition of instructions/advice (for advice aspect of consultation only).*	0	16 Duration and dominance: 0 Other: 16 *No differences in checking patient understanding, use of technical language, use of written materials, and various aspects of advice-giving (e.g. asking patient to repeat). Note that some variables did not appear at all in either modality.*	19 Duration and dominance: 2 Other: 17
Hammersley et al. (2019) [British general practice]	13 Duration and dominance: 1 Other: 12 *More problems discussed in face-to-face as compared to video (though not telephone) consultations. More patient education and counseling for both doctors and patients compared to both video and telephone, and more data gathering (except when compared to patients on the telephone). More direction-giving and partnership building compared to doctors in video consultations only.*	3 Duration: 0 Other: 3 *More rapport building (telephone for doctors and patients, video for patients only).*	26 Duration and dominance: 1 Other: 25 *No difference in number of problems when compared to telephone consultations. No differences in most RCGP indicators for either doctors or patients. On the RIAS, there were no differences in for doctors or patients in video or telephone for disagreement. All other RIAS variables showed at least one difference, although not always across every single comparison (see second column). Note that, for some comparisons, the variable did not appear at all in either modality.*	42 Duration and dominance: 2 Other: 40
Secondary care				
Frueh et al. (2007) [Therapy sessions]	0	1 *Competence rated higher when introducing "homework" assignments.*	18 *No differences in all other aspects of adherence to/competence with protocol, including therapist's explanations, empathy and rapport, and providing feedback.*	19

(*Continued*)

(Continued).

Paper	Comparisons favoring FtF	Comparisons favoring TM	No significant difference	Total no. of comparisons
Ball et al. (1995) [Psychiatry]	0	0	6 *No differences in non-verbal (angle of recline and mutual gaze) and verbal (interpretations, reflections, advisements, questions) behaviors.*	6
Day and Schneider (2002) [Psychotherapy]	0	2 *More client participation in both video and audio-only conditions.*	4 *No differences for client hostility or therapist explanation between face-to-face and video/audio-only.*	6
Demiris et al. (2005) [Dermatology	1 Duration and dominance: 1 Other: 0 *Longer.*	1 Duration and dominance: 0 Other: 1 *More small talk.*	6 Duration and dominance: 0 Other: 6 *No differences in assessment, psychosocial talk, education, treatment discussion, compliance promotion, and administrative talk.*	8 Duration and dominance: 1 Other: 7
Edison et al. (2013) [Dermatology]	1 Duration and dominance: 1 Other: 0 *More non-physician words (i.e. patients and companions spoke more).* *NOTE: A sub-variable of "Promoting compliance," "Positive reinforcement," was also significantly higher in face-to-face consultations.*	0	15 Duration and dominance: 1 Other: 14 *No significant difference in duration. Other main communication variables the same as Demiris et al., above. Edison et al. compared these variables twice (based on both number of words devoted to them and whether the variable was present), with no significant differences in either case.*	16 Duration and dominance: 2 Other: 14
Agha, Roter, et al. (2009) [Pulmonology]	3 Duration and dominance: 0 Other: 3 *Fewer requests for repetition and unintelligible utterances. Patients (but not doctors) give more orientation.*	0	26 Duration and dominance: 3 Other: 23 *No difference in duration, companion contribution, and dominance ratio. No differences in affect ratings, and no differences in all other RIAS variables (information exchange, types of talk, and rapport and partnership).*	29 Duration and dominance: 3 Other: 26
	39 Duration and dominance: 12 Other: 27	7 Duration and dominance: 0 Other: 7	143 Duration: 8 Other: 135	189 Duration: 20 Other: 169

[1]Refers to both variables that measured the length of the consultation, variables that measured professionals' dominance, and the closely related variables that were discussed alongside these in the main text (e.g. number of problems discussed).