Telemedicine in Primary Care: A Review of Quasi-Experimental Evidence

Viet Nguyen[1] and Mika Kortelainen[1,2]

Abstract

This review synthesizes evidence from thirteen quasi-experimental studies that evaluate the impact of telemedicine on primary care utilization, spending, and health outcomes across multiple countries. These studies span both pre-pandemic implementation and the accelerated adoption of remote care during COVID-19. Of these, five were classified as Tier 1, indicating a low risk of bias. Across studies, remote consultations most often complement rather than substitute for in-person visits, potentially contributing to higher overall utilization, particularly among new and low-frequency users. Evidence of substitution is limited and largely confined to specific contexts, such as mental health services or periods of restricted physical access during the pandemic. Health outcomes are generally comparable to in-person care, with emerging but still limited evidence of improvements in chronic disease management and mental health. Evidence on spending remains heterogeneous: costs tend to rise when telemedicine adds low acuity visits and follow up, but may be neutral or slightly lower when it replaces higher cost care.

Keywords: *Primary healthcare, telemedicine, effect, determinant, quasi-experiment*
JEL Classification: *I11, I12, I13, I18*

[1] University of Turku, Turku, 20014, Finland.
[2] Finnish Institute for Health and Welfare, Helsinki, 00271, Finland.
* Corresponding author(s). E-mail(s): tuanviet.t.nguyen@utu.fi
Contributing authors: mika.kortelainen@utu.fi

1 Introduction

In recent years, telemedicine services, such as video consultations, have rapidly gained traction among primary care providers across numerous countries. This shift has been significantly accelerated by the COVID-19 pandemic and the broader trend of ongoing digital transformation in healthcare systems [1, 2]. Telemedicine is commonly understood as the remote delivery of clinical services using information and communication technologies and is expected to improve health outcomes by enhancing access to care, particularly for individuals living in remote areas, vulnerable populations, and elderly people with chronic conditions [3, 4].

The scoping review by Erku et al. [5] highlights the importance of considering the readiness of the digital ecosystem, user behavior, broader health system requirements, and the capacity of primary healthcare when assessing the impact of digital health interventions. Similarly, a scoping review by Lakoma et al. [6] argues that tailoring digital solutions to meet the specific needs of different patient segments would increase their acceptability and suitability. Together, these studies emphasize the need for additional studies in primary care settings to further explore and understand the impacts of telemedicine services in real world contexts.

Although some systematic reviews [7–9] have explored the impacts of remote consultations, these existing reviews have largely been based on randomized controlled trials (RCT). Randomized controlled trials are widely regarded as the gold standard for studying causal relationships due to their ability to minimize bias through randomization. However, they are costly and time-intensive, resulting in only a limited number of studies featuring well-designed Class I RCT [10]. A recent systematic review by Campbell et al. analyzed 30 studies, of which only three were assessed as having a low risk of bias [9]. The review concluded that remote consultations may be as effective as face-to-face care and could improve efficiency and timeliness, while evidence on safety and patient centered outcomes remained inconclusive.

In this context, quasi-experimental designs offer a practical alternative to RCT to evaluate digital healthcare and telemedicine, particularly in assessing long-term and longitudinal effects, as well as broader outcomes. These designs allow for more flexibility in real-world settings where random assignment may not be feasible, while still providing valuable information on the effectiveness of interventions over time [10]. This review focuses on quasi-experimental studies and their methodologies for evaluating the impact of telemedicine in primary care settings. To our knowledge, this is the first review to systematically assess both findings and identification strategies in quasi-experimental telemedicine studies in primary care.

We conducted a narrative review rather than a systematic review due to resource constraints and the variation in methodological reporting among quasi-experimental studies. The broad and exploratory search strategy was better suited to this approach than the rigid criteria of a systematic review. Literature was identified through two sources: a search of the Embase database (see Annex for details) and relevant articles from the SoteDataLab project, a collaboration between the Finnish Institute for Health and Welfare (THL), Aalto University, and the University of Turku on digital healthcare and artificial intelligence. Additionally, study quality was assessed using the ROBINS-I tool [11]. According to Windle et al. [12], ROBINS-I represents the most comprehensive and granular tool available for assessing risk of bias in quasi-experimental studies.

2 Quasi-experimental studies for measuring impacts of telemedicine in primary care

2.1 Methodologies applied in the studies

Our review includes thirteen quasi-experimental studies evaluating the effects of telemedicine on utilization, spending, and clinical and quality-of-care outcomes (Table 1). These studies include six from the United States, four from Sweden, one from Australia, one from Israel, and one from Argentina. Relative to the COVID-19 pandemic (2019-2021), six

studies [13–18] were conducted pre-pandemic, focusing on routine remote care implementation and substitution patterns in standard healthcare settings. The remaining seven [19–25] were conducted during or after its onset, reflecting policy shifts and accelerated digital adoption.

The difference-in-differences (DiD) methods were the predominant approach, applied in eight studies [13, 14, 18, 19, 23–25]. Among these, several incorporated complementary techniques, such as matching and fixed effects (FE), to strengthen identification and control for unobserved heterogeneity. Notably, Avdic et al. [19] and Zeltzer et al. [25] exploit provider adoption to generate plausibly exogenous treatment exposure, thereby mitigating concerns about unobserved selection bias in observational settings. Dahlgren et al. [16] and Busso et al. [20] employ interrupted time series (ITS) and event study designs, respectively, which leverage pre-intervention dynamics and comparison groups and are therefore closely related to DiD-type approaches.

Instrumental variable (IV) methods were explicitly implemented in five studies [14, 17, 21–23], primarily to address endogeneity and selection bias arising from non-random exposure to telemedicine. These studies employed a range of instruments, including physician adoption of e-visits [14], local variations in COVID-19 exposure and shelter-in-place policies [21], cost-sharing waiver [23], and nurses' propensity to direct patients online [22]. While Bavafa et al. [14] used IV primarily as a robustness check, others [21–23] relied on their instruments as a central identification strategy to measure causal inference. In addition, Ellegård et al. [17] introduced an innovative fuzzy difference-in-discontinuities (Fuzzy Diff-in-Disc) design, which combines features of regression discontinuity and DiD frameworks. Leveraging an age-based user fee threshold as a natural experiment, this approach exploits both cross-sectional and temporal discontinuities to isolate the causal effects of direct-to-consumer telemedicine (DTC) pricing on healthcare utilization.

Taking a different approach, Dahlstrand [15] applied a value-added (VA) approach to estimate the causal impact of individual physicians on patient outcomes, leveraging the quasi-random assignment of patients to doctors in Sweden's online primary care system. The model controlled for patient characteristics, used only past data to estimate physician skill, adjusted for statistical noise and performance drift, and evaluated doctor effectiveness across multiple clinical outcomes. These estimates were subsequently used in a counterfactual optimization exercise that simulates positive assortative matching between patient risk and physician skill, quantifying the aggregate outcome gains achievable through reallocation of existing resources.

Table 1: *Summary of quasi-experimental studies on telemedicine in primary care*

Author (Year)	Setting	Method	Data Source	Primary Outcomes	Key Findings
Avdic et al. (2026) [19]	Nationwide telemedicine rollout among GPs	DiD with propensity adoption FE	Australian ABS PLIDA linked MBS– PBS individual-level data (2017–2022)	Prescribing quality	More intensive telemedicine adopters prescribed fewer antibiotics while maintaining prescribing quality.
Ashwood et al. (2017) [13]	DTC telehealth for acute respiratory infections	Matched DiD	US commercial claims data (2011–2013)	Total utilization, spending	DTC telehealth increased access but did not reduce spending. Most visits represented new utilization rather than substitution.
Bavafa et al. (2018) [14]	Active FFS patient e-visit adoption	DiD with FE (robustness checks: matched DiD, IV)	US primary care system data (2008–2013)	Visit frequency, HbA1c, LDL	E-visits increased office visits by about 6%. Effects on phone visits and patient health were mixed. Physicians accepted 15% fewer new patients after adoption.
Busso et al. (2021) [20]	Telemedicine during COVID-19	Event study	Argentina, major telemedicine provider (2019–2020)	Telemedicine demand (log daily calls)	Telemedicine demand rose sharply. First-time users increased, along with prescriptions, follow-up visits, and referrals.
Cantor and Whaley [21] (2021)	Telehealth response to reduced in-person visits	IV (SIP policy)	US proprietary telemedicine data (2016–2020)	Telehealth calls, substitution elasticity	Telemedicine partially substituted for in-person visits, with stronger effects among frequent users.
Dahlgren et al. (2023) [16]	First DTC telemedicine consultation	Matched ITS	Swedish registry data (Stockholm, 2016–2018)	Follow-up physician consultations	DTC telemedicine users increased healthcare utilization, mainly in the short term but also after 2–6 months. Total primary care consultations increased.
Dahlstrand (2025) [15]	Doctor–patient skill matching	Value-added approach	Swedish digital primary care and registry data (2016–2018)	ED visits, inappro priate antibiotic prescribing, follow-up visits	Matching high-risk patients with highly skilled doctors reduced ED visits by 4.4% and inappropriate antibiotic prescribing by 3.1%, while reducing income-based health inequality.

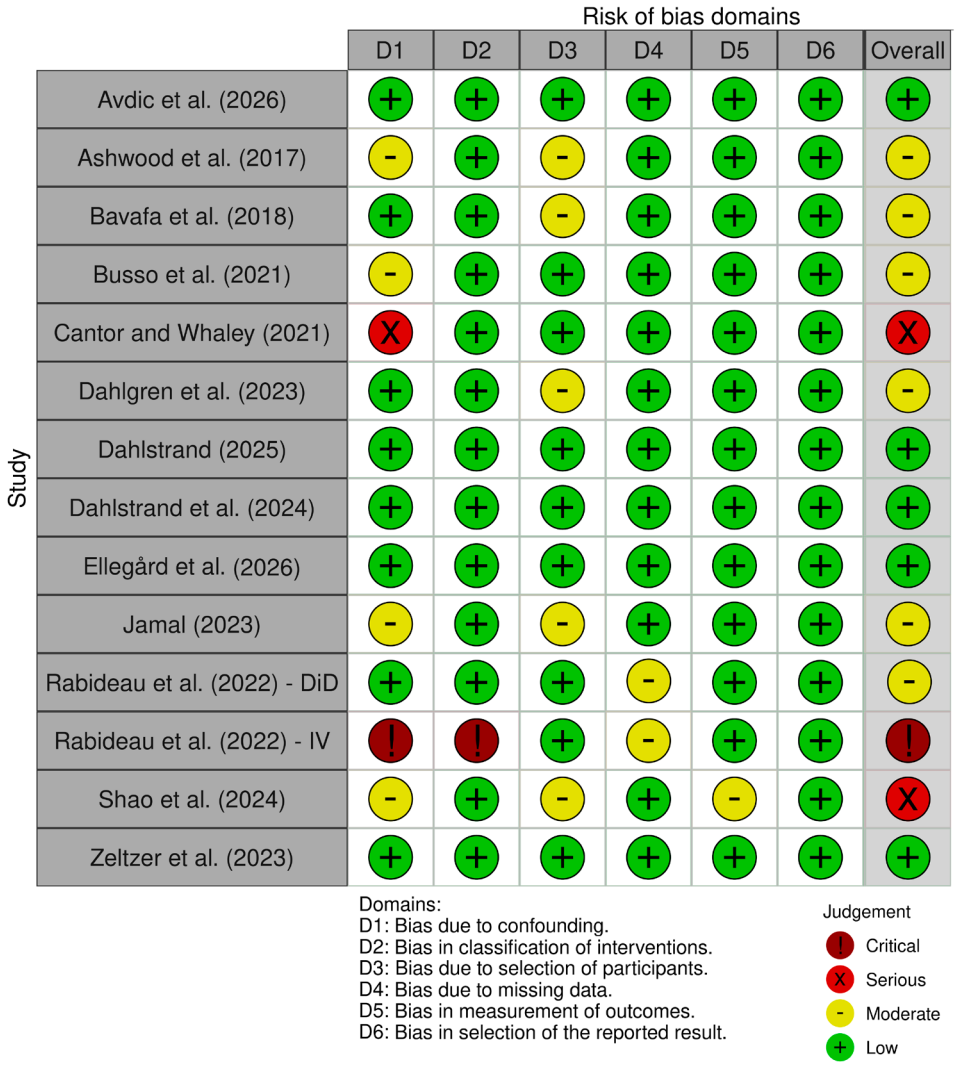
Author (Year)	Setting	Method	Data Source	Primary Outcomes	Key Findings
Dahlstrand et al. (2024) [22]	Random assignment to consultations	IV (nurse propensity)	Swedish primary care registries (2019–2020)	Consultation duration, hospitalization, ED visits, provider and patient costs	Online visits were faster and shorter. Out comes and patient satisfaction were similar. Short-term follow-up visits increased ED and in-person visits, with no medium-term health effects.
Ellegård et al. [17] (2026)	Young adults facing telemedicine price discontinuities	Fuzzy Diff-in-Disc (IV: age-20 fee cutoff)	Swedish registries (2012–2018)	Substitution rate of in-person visits, costs	Demand declined after user fees were introduced. Approximately 45% of online visits substituted for in-person care. Substitution was highest for respiratory infections.
Jamal (2023) [18]	Privately insured patients with chronic mental illness	Matched DiD	US employer-sponsored claims data (2009–2018)	Visits, spending, ED visits, hospitalization	Total visits increased while in-person visits decreased. No significant effects were found for ED visits or hospitalizations.
Rabideau and Eisenberg [23] (2022)	Telemedicine users under cost-sharing waivers (mental health)	DiD and IV	US employer-sponsored claims data (2018–2020)	Utilization, prescriptions, total spending	Cost-sharing waivers increased telemedicine use. Telemedicine substituted for in-person care, while inpatient and medication use declined.
Shao et al. (2024) [24]	Telehealth for type 2 diabetes	Matched DiD	US REACHnet EHR data (2020–2021)	HbA1c, LDL, BMI, BP	Telehealth reduced HbA1c levels. LDL, BMI, and blood pressure also improved.
Zeltzer et al. (2023) [25]	Telemedicine expansion	DiD with propensity adoption	Israeli Clalit Health Services data (2019– 2020)	Diagnostic accuracy, costs, ED use	Telemedicine slightly increased visits, reduced spending, and maintained care quality.

Notes: DTC = Direct-to-Consumer; DiD = Difference-in-Differences; FE = Fixed Effects; IV = Instrumental Variables; ITS = Interrupted Time Series; Diff-in-Disc = Difference-in-Discontinuities; FFS = Fee-for-Service; ABS PLIDA = Australian Bureau of Statistics Person-Level Integrated Data Asset; MBS = Medicare Benefits Schedule; PBS = Pharmaceutical Benefits Scheme; SIP = Shelter-in-Place; ED = Emergency Department; EHR = Electronic Health Record; BMI = Body Mass Index; BP = Blood Pressure; HbA1c = Glycated Hemoglobin; LDL = Low-Density Lipoprotein; COVID-19 = Coronavirus Disease 2019.

2.2 Quality assessment and causal identification

As mentioned in the Introduction, we assessed risk of bias using the latest version of ROBINS-I, released on 20 November 2025, which comprises six domains [11]. The results of the risk-of-bias assessment are presented in Figure 1. We grouped the thirteen studies into two tiers, while excluding the component of one study assessed as having a critical risk of bias from the primary conclusions.

Fig. 1: *ROBINS-I risk of bias assessment*



Tier 1 designs (low risk of bias):

Five studies are categorized as Tier 1, as they exploit plausibly exogenous variation or quasi-random assignment mechanisms that yield credible counterfactuals. These approaches substantially mitigate concerns related to unobserved selection bias. Specifically, Ellegård et al. [17] implement a fuzzy Diff-in-Disc design around an age-20 user-fee threshold, leveraging a sharp policy cutoff to identify price effects. Dahlstrand [15] exploits quasi-random patient allocation through a first-come, first-served mechanism on a digital platform. Zeltzer et al. [25] and Avdic et al. [19] base identification on predetermined supply-side exposure, using physicians' propensity to adopt telemedicine while conditioning on fixed

patient–provider relationships. Finally, Dahlstrand et al. [22] employ an instrumental variable strategy based on the quasi-random assignment of triage nurses, whose propensity to direct patients to digital care serves as a plausibly exogenous instrument, supported by institutional rules and balance diagnostics.

Tier 2 designs (moderate to serious risk of bias):

The remaining studies are grouped into Tier 2, encompassing those with moderate to serious risk of bias. Residual concerns primarily relate to confounding (Domain 1), participant selection (Domain 3). These studies typically rely on matched DiD. While these approaches are well suited to evaluating longitudinal effects and often incorporate rigorous matching procedures (e.g., propensity score or machine learning–based methods such as random forests) to improve balance on observed covariates, they remain susceptible to residual endogeneity arising from unobserved confounders and potential violations of identifying assumptions.

Rabideau and Eisenberg [23] comprises two distinct components, DiD and IV approaches. The IV component is assessed as having a critical risk of bias and accordingly, only the DiD estimates are retained for the conclusions.

2.3 Defining treatment and measuring outcomes

Evaluating quasi-experimental evidence in primary care requires careful attention to how both telemedicine exposure and outcomes are operationalized, because measurement choices can introduce endogeneity and obscure whether telemedicine substitutes for, or adds to, in-person care.

Across the reviewed studies, treatment is defined either as realized use or as potential access. Studies with moderate risk of bias often define adoption as the patient's first observed telemedicine visit [13, 16, 18, 24], an intuitive measure that is nonetheless endogenous to unobserved need, preferences, and digital capability. "Tier 1" designs instead anchor treatment in potential access or supply-side exposure. Zeltzer et al. [25] and Avdic et al. [19] classify patients using their pre-period affiliated physician's propensity to adopt telemedicine, holding patient–provider affiliations fixed to a pre-intervention period so that access is predetermined. Dahlstrand [15] similarly strengthens identification by exploiting quasi-random assignment on a digital platform, where a first-come-first-served mechanism conditionally allocates patients to doctors, eliminating endogenous sorting without requiring an external instrument.

A related measurement concern arises when exposure is defined as the share of services delivered virtually (Rabideau and Eisenberg [23], IV specification). This measure is endogenous because it reflects both the intervention and its own effects on physical service use, making any causal conclusions drawn from it invalid. This issue is largely avoided in designs using adoption timing or binary indicators rather than outcome-dependent shares.

Outcome definitions similarly affect interpretation. To address "consultation inflation", where telemedicine increases follow-ups rather than replacing in-person care, the most informative studies shift from visit counts to episode-based outcomes, grouping all contacts within a fixed post-index window (typically 7-30 days) and applying washout periods (often 14-30 days without prior use) to ensure the index visit reflects a new need [19, 22, 25]. This approach separates extensive-margin changes (more episodes) from intensive-margin changes (more contacts within an episode).

2.4 Findings from the studies

Utilization effects: Evidence from United States (US), Sweden and Israel shows that telemedicine often complements in-person care, leading to higher overall utilization among new and low-frequency patients [13, 14, 16, 17, 20, 22, 25]. In the US, Ashwood et al. [13] found that only 11.8% of DTC telehealth visits substituted for in-person care, with the majority representing new utilization for acute conditions. Bavafa et al. [14] reported similar patterns among active, fee-for-service patients with at least three visits per year, where telehealth adoption raised total utilization across routine

primary care encounters. In Sweden, research demonstrates consistent complementary effects in new or low-frequency patients. Ellegård et al. [17] found that about 45% of online consultations replaced in-person visits, while the remaining 55% added to total utilization. Dahlgren et al. [16] observed that online users had higher total service use than matched controls, with short-term increases in the first month and persistent effects up to 2–6 months post-consultation. Dahlstrand et al. [22] also report a greater likelihood of a new visit to a primary care provider within 30 days compared to in-person consultations. In Israel, Zeltzer et al. [25] show that telemedicine adoption similarly raises total service use, with higher short-term follow-up visits and overall engagement among new patients.

Evidence of substitution is more limited and generally context-dependent, occurring primarily in mental health services, through financial incentives such as cost-sharing waivers, and during external crises like the COVID-19 pandemic that restrict physical access. Cantor and Whaley [21] found that during the pandemic, telemedicine adoption coincided with a reduction in in-person visits. Similarly, Jamal [18] observed partial substitution in mental health care, where telemedicine increased total outpatient engagement while reducing in-person visits. Rabideau and Eisenberg [23] findings from the DiD approach also reported that cost-sharing waivers during COVID-19 increased telemedicine use, with telemedicine acting as a near-perfect substitute for in-person care.

Clinical and Quality-of-care Effects: Health effects vary by clinical and quality-of-care domain. In chronic disease management, telehealth has been associated with measurable improvements. Shao et al. [24] found improved glycemic control and better LDL and BMI outcomes among patients with type 2 diabetes, while matched primary care samples showed reduced rates of unhealthy HbA1c readings. Mental health care demonstrates similar stability or improvement. Jamal [18] reported increased outpatient mental health visits with no rise in ED use or hospitalization. Consistent with these patterns, Bavafa et al. [14] documented small but meaningful improvements in biometric indicators among e-visit adopters, most notably a 7–18% reduction in unhealthy HbA1c levels, and found no evidence of reduced laboratory monitoring.

Findings for acute or general care are also comparable to those for in-person care. Avdic et al. [19] show that greater telemedicine adoption is associated with lower antibiotic prescribing for acute infections, without evidence of deteriorating prescribing quality. Dahlstrand [15] and Zeltzer et al. [25] show that carefully designed matching or access expansions can reduce ED visits, inappropriate antibiotics and avoidable admissions, whereas another study by Dahlstrand et al. [22] document similar quality of care across modalities but an increase in short term ED use.

Spending Effects: Cost effects reflect the underlying utilization patterns. When telemedicine generates new visits, total spending tends to rise. Ashwood et al. [13] estimated a net annual spending increase of $45 per DTC telehealth user due to the predominance of new demand. While the unit cost of an online consultation is significantly lower than an in-person visit, this advantage is largely eliminated by subsequent utilization in hybrid settings [22]. When accounting for the higher incidence of short-term follow-ups, including primary care and ED visits, the provider's cost advantage for online consultations was drastically reduced to only 20% cheaper than in-person care. Furthermore, including follow-up costs meant that potential savings for patients were effectively eliminated, making online consultations approximately 6% more expensive in total compared to in-person care [22]. This cost inflation was pronounced for vulnerable patients (those with prior hospitalizations or ED visits), for whom online consultations increased total provider costs by 22% and patient costs by 64%.

Conversely, in specialized contexts like chronic care or during the COVID-19 lockdown period, overall spending remained neutral or decreased. For patients with chronic mental illness, the adoption of telemedicine resulted in no significant differences in overall medical spending or drug spending [18]. Zeltzer et al. [25] found that expanding telemedicine access in an integrated Israeli HMO after the COVID-19 lockdown generally reduced overall spending.

Separating "Tier 1" (low risk of bias) from the remaining studies does not materially change the overall interpretation. Tier 1 studies still support the same central conclusions: (i) telemedicine frequently complements in-person care and increases short-term downstream utilization [17, 22, 25]; (ii) quality of care appears broadly comparable between telemedicine and in-person care, with several studies documenting improvements in acute or general care outcomes. Greater

telemedicine adoption is associated with lower antibiotic prescribing and fewer ED visits in some settings [15, 19, 25]. However, there is still evidence of short term increases in ED use and follow up in person visits after telemedicine consultations [22]; and (iii) spending effects are context-dependent: episode costs may rise when telemedicine triggers additional follow-up [22], but may fall where it substitutes for higher-cost care [25].

3 Discussion

Evaluating the causal impact of telemedicine in primary care remains complex. Although modalities such as chat-based visits and video consultations, are increasingly integrated into primary care, isolating their independent effects on utilization, costs, and health outcomes is challenging. Randomized controlled trials, while the gold standard, face limitations in this context as mentioned in the introduction. In primary care, relevant outcomes such as mortality or quality of life are delayed or difficult to measure, while proxies like visit counts or follow-ups may reflect behavioral adjustments rather than clinical effectiveness [15]. Disparities in telemedicine adoption amplify these challenges. Evidence from the USA shows that older adults, racial minorities, and non-English speakers are less likely to use telemedicine [26–28], whereas in Sweden, younger, urban, and more educated individuals are more likely to engage [29–32]. Similar findings from China indicate that digital literacy, trust, and support systems strongly influence adoption [33].

Quasi-experimental designs provide a practical alternative, and the evidence offers valuable understanding of telemedicine's effects on productivity in healthcare and welfare systems. Across settings, the productivity implications depend largely on whether remote care substitutes for, or adds to, in-person services. When telemedicine primarily generates additional low-acuity consultations, especially among new or low-frequency users, overall productivity may fall, as providers spend more time delivering care without corresponding reductions in avoidable events or downstream costs. This pattern is visible in seven studies where virtual visits often complement rather than replace physical care [13, 14, 16, 17, 20, 22, 25]. Evidence of substitution remains limited to specific contexts such as mental health services, periods of restricted physical access during the pandemic, or when financial incentives like cost-sharing waivers are implemented [18, 21, 23]. This suggests that while telemedicine expands access, it may lead to an overall increase in service volume rather than a more efficient redistribution of existing care.

At the same time, telemedicine still has the potential to improve patient welfare by bridging access gaps. A systematic review by Capodici et al. [34] document this access channel through a significant weighted mean reduction of 25.4 days in waiting times. Our review complements this evidence by identifying settings where welfare gains arise from better targeting and care delivery, such as the Swedish context in which high-risk patients are effectively matched with skilled providers, with associated reductions in ED visits by 4.4% and inappropriate antibiotic prescribing by 3.1% [15]. As highlighted by De Guzman et al. [35], telemedicine is cost-effective when it reduces overall demand or redirects patients from high-cost secondary care to lower-cost primary care. However, our review also warns of "consultation inflation", whereby virtual contacts trigger additional follow-up care and may increase net system spending [22]. Productivity gains are therefore contingent on the efficient use of clinician time; if the lower unit cost of a virtual visit is offset by administrative burdens or subsequent in-person follow-ups, the economic advantage is diminished.

Consistent with the RCT-based review by Campbell et al. [9], our synthesis suggests that health outcomes following remote consultations are broadly comparable to those observed in in-person care. Quasi-experimental evidence adds more insight into specific domains. In chronic disease management, studies report emerging benefits, including improvements in glycaemic control (HbA1c) and BMI among patients with diabetes [14, 24]. In acute or general care, several studies indicate that quality of care is either improved or at least maintained [15, 19, 25], although another study documents increases in short-term ED visits after virtual consultations [22].

The "Tier 2" studies included in our review are subject to well-documented limitations related to residual confounding, endogeneity, and the incomplete observability of key determinants of treatment and outcomes. In particular, these designs often struggle to account for unobserved factors such as digital access and literacy, provider quality, patient preferences, family or community support, and concurrent policy changes, all of which may simultaneously influence telemedicine

adoption and health outcomes [13, 18, 20, 24]. Identification problems are especially acute in the study by Cantor and Whaley [21], where instruments may be correlated with broader contextual forces such as local pandemic intensity, thereby violating exclusion restrictions and blurring causal interpretation. Several studies are also vulnerable to missing data and outcome measurement limitations, for example when telemedicine exposure is undefined during periods of non-use or when clinical testing is disrupted [23, 24].

A central concern is selection bias, as patients and providers self-select into remote care based on factors such as digital literacy, health seeking behavior, baseline health status, or technology access. In econometric terms, this reflects selection into treatment and potential violations of the conditional independence assumption. Within the ROBINS-I framework, these concerns are captured under domain 1, which provides a structured approach to assessing whether observed and unobserved confounders are adequately addressed. At the same time, selection operates not only through treatment assignment but also through inclusion into the analytic sample. Many studies implicitly restrict their samples based on healthcare utilization or post-intervention behavior, thereby introducing additional selection concerns [13, 14, 16, 18, 24]. These issues are captured under domain 3 of ROBINS-I, which highlights the risk that the criteria selection may not be representative of the target population.

Difference-in-Differences studies that define treatment based on the first observed telemedicine use are typically rated as having moderate risk of bias in domain 3, yet they remain an important and defensible quasi-experimental approach. Under ROBINS-I, forming cohorts using post-intervention characteristics can induce selection bias by excluding individuals who never enter the treated pathway. However, studies such as Dahlgren et al. [16] and Jamal [18] address this concern by applying rigorous matching, from propensity score matching to random forest proximity matching, to balance treated and control groups on observable baseline covariates. When these designs also support the parallel trends assumption and pass falsification or placebo tests, they can yield credible causal estimates. Although other "Tier 1" studies achieve stronger internal validity through context-specific sources of exogenous variation, such as abrupt policy rollouts [17], physician adoption propensity [19, 25], or quasi-random patient–doctor assignment [15], such identification opportunities are not always available. In their absence, "first-telemedicine-use" designs provide a practical alternative for assessing the longitudinal effects of telemedicine in routine care, using broad populations and rich administrative data that randomized trials often cannot capture.

Beyond DiD-based approaches, a subset of studies employed IV strategies, with heterogeneous identification strength. As Windle et al. [12] note, ROBINS-I targets "follow-up studies" and is therefore less applicable to IV studies. This pattern is also evident in recent economics reviews, which have applied ROBINS-I mainly to DiD designs [36, 37]. Unlike DiD approaches, where key identifying assumptions such as parallel trends can be partially assessed through observable diagnostics, IV strategies rely on the exclusion restriction, which is inherently untestable.

Given these limitations, we consider IV studies separately for deeper evaluation of identification strength and risk of bias (see Appendix for details). Designs based on quasi random assignment or sharp policy thresholds, such as nurse assignment in Dahlstrand et al. [22] and age based user fee discontinuities in Ellegård et al. [17], provided the most credible identification, supported by strong first stage relationships and transparent falsification tests. Other IV approaches, including provider intensity instruments and pandemic related policy variation while often exhibiting strong first stages, raised greater concerns regarding exclusion restrictions and external validity [14, 21].

Another limitation is inconsistent definitions of "digital consultation", sometimes conflating video, chat, or platform-based consultations with traditional telephone encounters, add further complexity. In fact, while the major studies in this review clearly defined the types of remote consultations available on digital platforms, some lacked clarity regarding the inclusion of traditional telephone consultations [18, 23, 24].

Telemedicine may improve resource allocation by enabling remote consultations, reducing travel burdens, and extending provider capacity. Yet evidence of increased utilization following remote care raises questions about whether this reflects the fulfillment of previously unmet needs or the emergence of convenience-driven use. Similarly, while digital tools may enhance access for underserved or rural groups, persistent disparities in digital literacy, internet access, and affordability

continue to limit uptake. Given the still-limited evidence on chronic disease management and mixed effects on spending, future work should prioritize robust, large-scale studies that evaluate whether digital modalities can produce sustained improvements in health outcomes and cost efficiency for patients with long-term conditions.

4 Conclusions

This review synthesizes quasi-experimental evidence on the causal effects of telemedicine in primary care, extending the existing literature by incorporating broader populations, more diverse settings, and a wider range of outcome measures. Overall, health outcomes associated with remote consultations appear generally comparable to in-person care, with emerging, but still limited, signals of benefit in chronic disease management and mental health. However, virtual visits often complement rather than substitute for in-person care, raising concerns about potential convenience-driven overuse and uncertain effects on overall spending. At the same time, persistent disparities in digital literacy, trust, and access highlight ongoing inequities in who benefits from telemedicine in primary care. As telemedicine continues to expand, future work should prioritize rigorous designs that further strengthen causal inference, assess long-term effects, particularly for chronic disease care, and incorporate patient perspectives to ensure that telemedicine in primary care advances both efficiency and equity within healthcare and welfare systems.

References

[1] Berry CA, Kwok L, Massar R, Chang JE, Lindenfeld Z, Shelley DR, et al. Patients' Perspectives on the Shift to Telemedicine in Primary and Behav-ioral Health Care during the COVID-19 Pandemic. J Gen Intern Med. 2022;37(16):4248–4256. https://doi.org/10.1007/s11606-022-07827-4.

[2] OECD.: The future of telemedicine after COVID-19 [Web Page]. Available from: https://www.oecd.org/coronavirus/policy-responses/the-future-of-telemedicine-after-covid-19-d46e9a02/.

[3] Neculau AE, Lăcătus, AM, Mărginean M, Fărcăs, anu D. A guide to telemedicine in primary healthcare. UNICEF; 2022. Available from: https://www.unicef.org/romania/media/10891/file/Aguidetotelemedicineinprimaryhealthcare.pdf.

[4] World Health Organization. Regional digital health action plan for the WHO European Region 2023–2030. World Health Organization; 2022. Available from: https://iris.who.int/bitstream/handle/10665/360950/72wd05e-DigitalHealth-220529.pdf?sequence=2&isAllowed=y.

[5] Erku D, Khatri R, Endalamaw A, Wolka E, Nigatu F, Zewdie A, et al. Digital Health Interventions to Improve Access to and Quality of Primary Health Care Services: A Scoping Review. Int J Environ Res Public Health. 2023;20(19). https://doi.org/10.3390/ijerph20196854.

[6] Lakoma S, Pitkänen L, Lahdensuo K, Lillrank P, Torkki P. Digital primary care visits designed for different patient segments in the pre-pandemic era: A scoping review. Finnish Journal of eHealth and eWelfare. 2023;15(3):287–304. https://doi.org/10.23996/fjhw.125966.

[7] Mold F, de Lusignan S, Sheikh A, Majeed A, Wyatt JC, Quinn T, et al. Patients' online access to their electronic health records and linked online services: a sys-tematic review in primary care. Br J Gen Pract. 2015 March 1;65(632):e141–e151. https://doi.org/10.3399/bjgp15X683941.

[8] Albornoz CdS, Sia KL, Harris A. The effectiveness of teleconsultations in pri-mary care: systematic review. Fam Pract. 2022;39(1):168–182. https://doi.org/10.1093/fampra/cmab077.

[9] Campbell K, Greenfield G, Li E, et al. The Impact of Virtual Consultations on the Quality of Primary Care: Systematic Review. Journal of Medical Internet Research. 2023;25:e48920. https://doi.org/10.2196/48920.

[10] Chumbler NR, Kobb R, Brennan DM, Rabinowitz T. Recommendations for research design of telehealth studies. Telemed J E Health. 2008;14(9):986–9. https://doi.org/10.1089/tmj.2008.0108.

[11] Sterne J, Higgins J, Group RIVD.: ROBINS-I V2: Risk Of Bias In Non-randomized Studies of Interventions. Accessed: 02 December 2025. https://www.riskofbias.info/welcome/robins-i-v2.

[12] Windle SB, Harper S, Arneja J, Socha P, Nandi A. Systematic Reviews of Quasi-Experimental Studies: Challenges and Considerations. Journal of Clinical Epidemiology. 2026; https://doi.org/10.1016/j.jclinepi.2025.112121.

[13] Ashwood JS, Mehrotra A, Cowling D, Uscher-Pines L. Direct-To-Consumer Tele-health May Increase Access To Care But Does Not Decrease Spending. Health Aff (Millwood). 2017;36(3):485–491. https://doi.org/10.1377/hlthaff.2016.1130.

[14] Bavafa H, Hitt LM, Terwiesch C. The Impact of E-Visits on Visit Frequencies and Patient Health: Evidence from Primary Care. Manage Sci. 2018;64(12):5461–5480. https://doi.org/10.1287/mnsc.2017.2900.

[15] Dahlstrand A. Defying distance? The provision of medical services in the digital age. No. 70 in URPP Equality of Opportunity Discussion Paper Series; 2025. Available from: https://www.urpp-equality.uzh.ch/dam/jcr:c401fa00-b718-4f35-8047-1134420b8183/70 Defying Distance 250707.pdf.

[16] Dahlgren C, Spanberg E, Svereus S, Dackehag M, Wandell P, Rehnberg C. Short-and intermediate-term impact of DTC telemedicine consultations on subsequent healthcare consumption. Eur J Health Econ. 2024;25(1):157–176. https://doi.org/10.1007/s10198-023-01572-z.

[17] Ellegård LM, Kjellsson G, Mattisson L. An App Call a Day Keeps the Patient Away? Substitution of Online and In-Person Doctor Consultations Among Young Adults. The Economic Journal. 2026 01;p. ueag016. https://doi.org/10.1093/ej/ueag016. https://academic.oup.com/ej/advance-article-pdf/doi/10.1093/ej/ueag016/66631292/ueag016.pdf.

[18] Jamal A. Effect of Telemedicine Use on Medical Spending and Health Care Utilization: A Machine Learning Approach. AJPM Focus. 2023;2(3):100127. https://doi.org/10.1016/j.focus.2023.100127.

[19] Avdic D, Kunz JS, M´endez SJ, Wi´sniewska M. Does telemedicine technol-ogy affect prescribing quality in primary care? The case of antibiotics. Journal of Health Economics. 2026;105:103096. https://doi.org/10.1016/j.jhealeco.2025.103096.

[20] Busso M, Gonzalez MP, Scartascini C. On the demand for telemedicine: Evidence from the COVID-19 pandemic. Health Economics. 2022;31(7):1491–1505. https://doi.org/10.1002/hec.4523.

[21] Cantor JH, Whaley CM. The Doctor Will See You: Online Physician Transforma-tions During COVID-19. Santa Monica, CA: RAND Corporation; 2021. Available from: https://www.rand.org/pubs/working papers/WRA621-5.html.

[22] Dahlstrand A, Nestour NL, Michaels G. Online versus In-Person Services: Effects on Patients and Providers. Bonn: Institute of Labor Economics (IZA); 2024. 17177. Available from: https://www.iza.org/publications/dp17177.

[23] Rabideau B, Eisenberg MD. The Effects of Telemedicine on the Treatment of Mental Illness: Evidence from Changes in Health Plan Benefits. 2022 mar;SSRN Working Paper. https://doi.org/10.2139/ssrn.4065120.

[24] Shao Y, Shi L, Nauman E, Price-Haywood E, Stoecker C. Telehealth use and its impact on clinical outcomes in patients with type 2 diabetes during the COVID-19 pandemic. Diabetes Obes Metab. 2024;26(1):118–125. https://doi.org/10.1111/dom.15293.

[25] Zeltzer D, Einav L, Rashba J, Balicer RD. The Impact of Increased Access to Telemedicine. Journal of the European Economic Asso-ciation. 2024 04;22(2):712–750. https://doi.org/10.1093/jeea/jvad035. https://academic.oup.com/jeea/article-pdf/22/2/712/57146916/jvad035.pdf.

[26] Reed ME, Huang J, Graetz I, Lee C, Muelly E, Kennedy C, et al. Patient Characteristics Associated With Choosing a Telemedicine Visit vs Office Visit With the Same Primary Care Clinicians. JAMA Netw Open. 2020;3(6):e205873. https://doi.org/10.1001/jamanetworkopen.2020.5873.

[27] Eberly LA, Kallan MJ, Julien HM, et al. Patient characteristics associated with telemedicine access for primary and specialty ambulatory care during the COVID-19 pandemic. JAMA Network Open. 2020;3(12):e2031640–e2031640. Published 2020 Dec 1, published correction appears in JAMA Netw Open. 2021 Feb 1;4(2):e211913. doi: 10.1001/jamanetworkopen.2021.1913. https://doi.org/ 10.1001/jamanetworkopen.2020.31640.

[28] Ritzwoller DP, Goodrich GW, Tavel HM, Odelberg MR, Davis TL, Gander JC, et al. Patient Factors Associated With Use of Adult Primary Care and Virtual Visits During the COVID-19 Pandemic. Med Care. 2023;61(Suppl 1):S12–s20. https://doi.org/10.1097/mlr.0000000000001792.

[29] Ekman B, Thulesius H, Wilkens J, Lindgren A, Cronberg O, Arvidsson E. Utilization of digital primary care in Sweden: Descriptive analysis of claims data on demographics, socioeconomics, and diagnoses. Int J Med Inform. 2019;127:134–140. https://doi.org/10.1016/j.ijmedinf.2019.04.016.

[30] Dahlgren C, Dackehag M, Wandell P, Rehnberg C. Determinants for use of direct-to-consumer telemedicine consultations in primary healthcare-a reg-istry based total population study from Stockholm, Sweden. BMC Fam Pract. 2021;22(1):133. https://doi.org/10.1186/s12875-021-01481-1.

[31] Dahlstrand A, Farrokhnia N. Socioeconomic, medical and demographic char-acteristics of early adopters of digital primary care. Scandinavian jour-nal of public health. 2023;p. 14034948221119640. https://doi.org/10.1177/14034948221119640.

[32] Söderberg D, Bonn SE, Sjöblom L, Dahlgren A, Muli I, Amer-Wåhlin I, et al. Individual Patient Factors Associated with the Use of Physical or Digital Pri-mary Care in Sweden. Telemedicine journal and e-health : the official journal of the American Telemedicine Association. 2024; https://doi.org/10.1089/tmj.2023.0351.

[33] Du Y, Zhou Q, Cheng W, Zhang Z, Hoelzer S, Liang Y, et al. Factors Influ-encing Adoption and Use of Telemedicine Services in Rural Areas of China: Mixed Methods Study. JMIR Public Health Surveill. 2022;8(12):e40771. https://doi.org/10.2196/40771.

[34] Capodici A, Noci F, Nuti S, Emdin M, Dalmiani S, Passino C, et al. Reducing outpatient wait times through telemedicine: a systematic review and quantitative analysis. BMJ Open. 2025;15(1). https://doi.org/10.1136/bmjopen-2024-088153. https://bmjopen.bmj.com/content/15/1/e088153.full.pdf.

[35] De Guzman KR, Snoswell CL, Caffery LJ, Smith AC. Economic evaluations of videoconference and telephone consultations in primary care: A systematic review. Journal of telemedicine and telecare. 2024;30(1):3–17. https://doi.org/10.1177/1357633X211043380.

[36] Shawa KC, Hollingsworth B, Zucchelli E. A systematic review and meta-analysis on the effects of ill health and health shocks on labour supply. Systematic Reviews. 2024;13:52. https://doi.org/10.1186/s13643-024-02454-y.

[37] Salguero RAF.: Effectiveness of Carbon Pricing and Compensation Instruments: An Umbrella Review of the Empirical Evidence. Available from: https://arxiv. org/abs/2512.06887.

Statements and Declarations

Mika Kortelainen (MK) conceived the idea for the article and provided essential reference materials. Viet Nguyen (VN) performed the search strategy, extracted the data, and drafted the manuscript. MK supervised the quality of the work and conducted a critical review of the study. All authors read and approved the final manuscript for submission.

Declarations

Funding
This study was funded by the Ministry of Social Affairs and Health (Finland), Strategic Research Council (grant no. 372591) and Eino Jutikkala fund. The funders had no role in the design, execution, interpretation, or writing of the study.

Competing interests
The authors declare no competing interests.

Ethics approval
Not applicable

Consent to participate
Not applicable

Consent for publication
Not applicable

Availability of data and material
Not applicable

Annex

Search terms:
Sources: Embase
Date: 26/04/2024
Results: 5,842
Details: ('telemedicine'/exp OR 'tele medicine':ti,ab,kw OR 'telemedicine':ti,ab,kw OR 'virtual medicine':ti,ab,kw OR 'telehealth'/exp OR 'e-health':ti,ab,kw OR 'ehealth':ti,ab,kw OR 'tele-health':ti,ab,kw OR 'telehealth':ti,ab,kw OR 'tele-consulta-tion'/exp OR 'long distance consultation':ti,ab,kw OR 'remote consultation':ti,ab,kw OR 'tele-consultation':ti,ab,kw OR 'teleconsultation':ti,ab,kw OR 'telephone consulta-tion':ti,ab,kw OR 'telephone-based consultation':ti,ab,kw OR 'digital health'/exp OR 'digital health':ti,ab,kw OR 'remote consultation' OR 'e health':ti,ab OR 'e counseling'/exp OR 'distance counseling':ti,ab,kw OR 'e-counseling':ti,ab,kw OR 'online counseling':ti,ab,kw OR 'remote counseling':ti,ab,kw OR 'tele counseling':ti,ab,kw OR 'video consultation'/exp OR 'telemedicine video-consultation':ti,ab,kw OR 'video consultation':ti,ab,kw OR 'videoconsul-tation':ti,ab,kw) AND ('counseling'/exp OR 'counseling':ti,ab,kw OR 'counselling':ti,ab,kw OR 'primary health care'/exp OR 'first line care':ti,ab,kw OR 'health care, primary':ti,ab,kw OR 'primary care nursing':ti,ab,kw OR 'primary health care':ti,ab,kw OR 'primary health-care':ti,ab,kw OR 'primary nursing care':ti,ab,kw OR 'general practitioner'/exp OR 'gp (general practitioner)':ti,ab,kw OR 'family doctor':ti,ab,kw OR 'family physician':ti,ab,kw OR 'family physicians':ti,ab,kw OR 'family practitioner':ti,ab,kw OR 'general physi-cian':ti,ab,kw OR 'general practice physician':ti,ab,kw OR 'general practitioner':ti,ab,kw OR 'general practitioners':ti,ab,kw OR 'physicians, family':ti,ab,kw OR 'physicians, pri-mary care':ti,ab,kw OR 'practitioner, general':ti,ab,kw OR 'primary care doctor':ti,ab,kw OR 'primary care physician':ti,ab,kw OR 'primary care physicians':ti,ab,kw OR 'primary medical care'/exp OR 'medical care, primary':ti,ab,kw OR 'primary care':ti,ab,kw OR 'primary medical care':ti,ab,kw) AND ('randomized controlled trial'/exp OR 'controlled trial, randomized':ti,ab,kw OR 'randomised controlled study':ti,ab,kw OR 'randomised con-trolled trial':ti,ab,kw OR 'randomized controlled study':ti,ab,kw OR 'randomized controlled trial':ti,ab,kw OR 'trial, randomized controlled':ti,ab,kw OR 'controlled study'/exp OR 'con-trol group study':ti,ab,kw OR 'control group trial':ti,ab,kw OR 'controlled study':ti,ab,kw OR 'controlled trial':ti,ab,kw OR 'placebo'/exp OR 'case control study'/exp OR 'case con-trol study':ti,ab,kw OR 'case-control studies':ti,ab,kw OR 'case-control study':ti,ab,kw OR 'control study, case':ti,ab,kw OR 'matched case control':ti,ab,kw OR 'matched case control studies':ti,ab,kw OR 'matched case control study':ti,ab,kw OR 'crossover procedure'/exp OR 'cross over clinical study':ti,ab,kw OR 'cross over clinical trial':ti,ab,kw OR 'cross over comparison':ti,ab,kw OR 'cross over design':ti,ab,kw OR 'cross over method':ti,ab,kw OR 'cross over procedure':ti,ab,kw OR 'cross over study':ti,ab,kw OR 'cross over test':ti,ab,kw OR 'cross over trial':ti,ab,kw OR 'cross-over studies':ti,ab,kw OR 'crossover clinical study':ti,ab,kw OR 'crossover clinical trial':ti,ab,kw OR 'crossover comparison':ti,ab,kw OR 'crossover design':ti,ab,kw OR 'crossover method':ti,ab,kw OR 'crossover procedure':ti,ab,kw OR 'crossover study':ti,ab,kw OR 'crossover test':ti,ab,kw OR 'crossover trial':ti,ab,kw OR 'double blind cross over study':ti,ab,kw OR 'double blind crossover study':ti,ab,kw OR 'procedure, crossover':ti,ab,kw OR 'observational study'/exp OR 'meta analysis'/exp OR 'analysis, meta':ti,ab,kw OR 'meta analysis':ti,ab,kw OR 'meta-analysis':ti,ab,kw OR 'metaanalysis':ti,ab,kw OR 'systematic review'/exp OR 'review, systematic':ti,ab,kw OR 'systematic review':ti,ab,kw)

Assessment on Instrumental Variable Studies

Study	Instrumental Variable (IV)	Supporting Tests & F-Statistics	Uncertainty / Concerns	Final Assessment
Dahlstrand et al. (2024)	Nurses' propensities to direct patients to online consultations	First-stage F > 100 balance tests on patient observables institutional rules (nurses cannot prescribe or refer)	Treatment bundled with doctor identities. Nurse "mistake" rates could proxy for skill, though tests find no correlation	**Strong.** Clean quasi-random assignment with limited exposure to missing-data or reporting bias. Patient selection mitigated by quasi-random nurse assignment and balance tests Exclusion supported because nurses cannot prescribe, refer, or order tests
Ellegård et al. (2026)	Age-based discontinuity at 20th birthday triggering user fees	First-stage F = 157.1 placebo cutoffs (ages 19, 21) donut RD to address anticipation	Possible age-20 confounders. intertemporal substitution around cutoff	**Strong.** Credible local identification with transparent assumptions. Sorting and anticipation tested via donut RD and bandwidth robustness Spurious age effects ruled out by placebo thresholds
Bavafa et al. (2018)	Physician Intensity of E-Visit Adoption: The number of e-visits a provider conducts with all other patients in a given month ("leave-one-out" instrument)	First-stage F = 135.36 auxiliary tests showing no change in visit duration or labs	Exclusion restriction untestable Physician time-allocation or technology adoption may affect outcomes directly Wide confidence intervals	**Moderately strong but assumption-dependent IV**
Cantor & Whaley (2021)	County-level exposure to SIP policies and timing of first COVID case/death	First-stage F = 21–39 event study showing post-policy demand break	Policy endogeneity Contemporaneous nationwide shocks Pandemic-era measurement issues	**Weak identification. Acceptable first stage but substantial threats to exclusion and interpretation**
Rabideau & Eisenberg (2022)	Cost-sharing waiver	F ≈ 2111 re-estimation holding total service quantity fixed	The IV analysis defines treatment intensity as the quarterly share of services delivered virtually, raising concerns regarding classification and denominator endogeneity.	**These issues indicate a critical risk of bias, and the related findings were therefore excluded.**