



**HKU  
Med**

School of Clinical Medicine  
Department of Family Medicine  
& Primary Care  
香港大學家庭醫學及基層醫療學系

REIMAGINING MEDICAL TRAINING WITH AI

# Preparing Medical Students for Digital Care

Piloting a GenAI-Supported Longitudinal, Outcomes-Based Telemedicine Curriculum in Family Medicine (Interim Results)

## Dr. Diana Wu

MD (Canada), FCFP (Canada), FHEA  
Lecturer

Department of Family Medicine and Primary Care  
Medical Ethics and Humanities Unit  
School of Clinical Medicine, HKU

## **01** Telemedicine in context

Global trend, the HK landscape, and current curriculum gaps

## **02** The pilot curriculum

Outcomes-based design, integrated across Y4 and Y6 Family Medicine

## **03** GenAI-supported feedback

Workflow, worked examples, and student response

## **04** Interim evaluation

E-modules, GenAI feedback, and self-efficacy results

## **05** Key takeaways

What's working, what's next



# Telemedicine: definition and the HK context

## DEFINITION

Remote interactions between patients and providers — using technology to enhance care quality.

## MODES OF DELIVERY

- Video consultation
- Telephone consultation
- Secure messaging

## IN HONG KONG

### **HA Go delivers remote care across the public sector**

Telemedicine is scaling rapidly across HK (particularly in private sector) and the GBA region — but undergraduate medical training has not kept pace.

## TRAINING GAP

*Students enter clinical years without structured teleconsultation training*

# Competency frameworks: limited consensus, especially in Asia

*Two frameworks anchor international thinking on telehealth competencies. None exist currently for the Asian undergraduate context.*

## Association of American Medical Colleges (AAMC)

### Telehealth Competencies Across the Learning Continuum

**MENU** **AAMC**

### Telehealth Competencies

The competencies are organized across six domains:

1. Patient Safety and Appropriate Use of Telehealth
2. Access and Equity in Telehealth
3. Communication via Telehealth
4. Data Collection and Assessment via Telehealth
5. Technology for Telehealth
6. Ethical Practices and Legal Requirements for Telehealth

## CanMEDS 2025

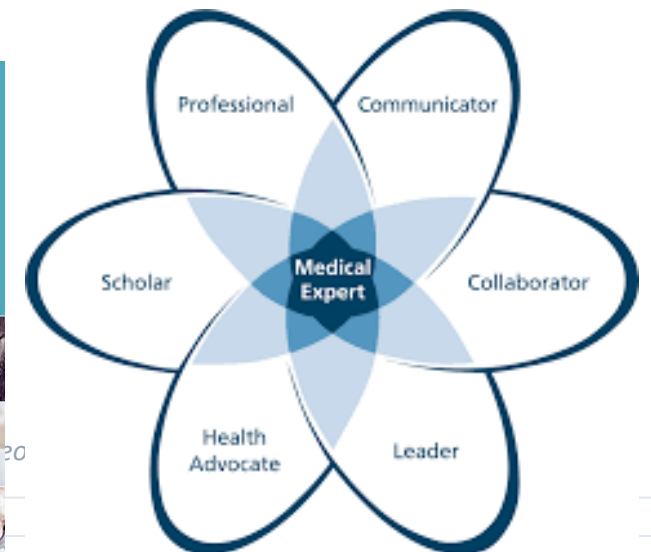
### Virtual Care Competencies

**VIRTUAL CARE  
IN CANADA:  
PROGRESS AND POTENTIAL**

REPORT OF THE VIRTUAL CARE  
TASK FORCE

FEBRUARY 2022

Logos: Association of Medical Colleges of Canada, Canadian Medical Association, Faculty of Medicine of the University of Toronto, University of Ottawa, University of Saskatchewan, Royal College of Physicians and Surgeons of Canada, Royal College of Physicians and Surgeons of the United Kingdom, Royal College of Physicians and Surgeons of the United Kingdom, Royal College of Physicians and Surgeons of the United Kingdom.





# Existing HKU FM curriculum: fragmented and lack integrated framework

**01**

## Zoom patient interactions only

Not purpose-built (“history-taking practice” since COVID)

No accompanying telemedicine curriculum

**02**

## Limited structured feedback

Tutor capacity doesn't scale to the volume of practice consultations.

Feedback is given on **ONE** patient interaction during a Video Review session with a tutor

**03**

## No alignment with competency standards

Telemedicine skills not formally assessed nor aligned with international standards.



# Pilot objectives

*Develop an outcomes-based, evidence-informed, longitudinal telemedicine curriculum that...*

**A**

## **Aligns to international competency frameworks**

Maps to AAMC and CanMEDS 2025 telehealth competencies – foundational knowledge, “web-side” skills, ethics/legal aspects

**B**

## **Targets feasibility, feedback quality, and acceptance**

Evaluation follows Kirkpatrick's four-level model.

**C**

## **Delivers personalised, scalable feedback**

GenAI-generated, rubric-aligned, tutor-oversight.

**D**

## **Longitudinal & operates without additional manpower**

Runs within existing Y4 and Y6 FM clerkship structures.



# Participants

YEAR 4 · GENERAL CLERKSHIP

# 4

teleconsultation  
sessions per student **over 1 day**

COHORT

**288 Y4 students**

FORMAT

6-min Zoom consultations with standardized  
patient actors

Done on site (Health Systems)

YEAR 6 · ASSISTANT INTERNSHIP

# 12

teleconsultation  
sessions per student **over 2 days**

COHORT

**301 Y6 students**

FORMAT

10-min Zoom consultations with standardized  
patient actors

Done at home



# Curriculum: four integrated components

*E-modules come first, group video review last; ethics workshop and teleconsultations run in parallel and can be completed in either order.*

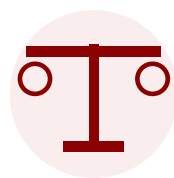


ASYNCHRONOUS

## E-modules

- Introduction to Telemedicine
- History-taking in Telemedicine

2 modules · self-paced



SYNCHRONOUS

## Ethics workshop

- Consent & MCHK guidance
- Conflict of interest
- Equity, bias, escalation & safety

Live · facilitated discussion



PRACTICE

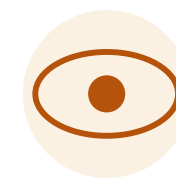
## Teleconsultations

- Zoom with standardized patients

Y4  
4 sessions

Y6  
12 sessions

GEN-AI FEEDBACK



REFLECTION

## Group video review

- Tele-OSCE checklist
- Peer + tutor debrief

Small group · shared learning



## GenAI Feedback responds to existing student feedback:

### GAP 1

Students want individualized feedback on all online practice consultations (vs just 1)

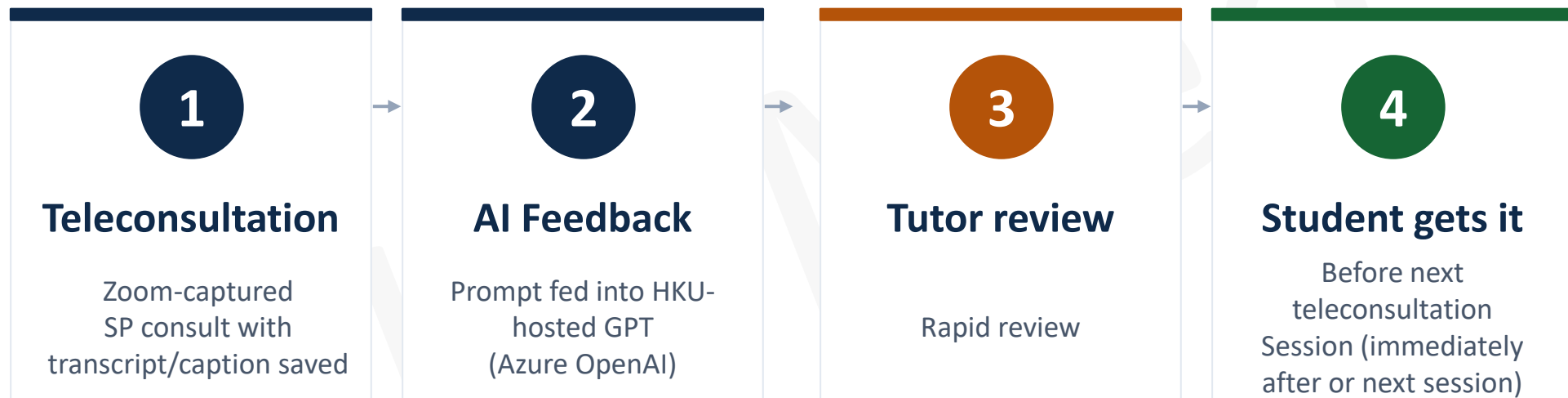
### GAP 2

Limited tutor capacity to deliver timely feedback at scale (1 tutor: 6 students)

# GenAI feedback workflow



*Four-step pipeline — delivers feedback before the next teleconsultation, with tutor oversight throughout.*



## KEY DESIGN PRINCIPLE

*GenAI fills the feedback bottleneck — formative feedback only*



RESULTS

# E-modules

**Telemedicine: An Introduction**

START COURSE

This interactive course introduces senior medical students in Hong Kong to the delivery of healthcare through virtual platforms. Through two structured e-learning modules followed by practical session and an in-person ethics workshop, students will explore the foundations, practical applications, and ethical

Course Preview Current version

Telemedicine: An Introduction  
0% COMPLETE

Section 1: Current Telemedicine Landscape

Section 2: In-person VS telemedicine encounters

Section 3: Practical Issues

**HKU Med** School of Clinical Medicine  
Department of Family Medicine & Primary Care  
香港大學家庭醫學及基層醫療學系

**FM Telemedicine Platform**

Continue with HKU Portal

or

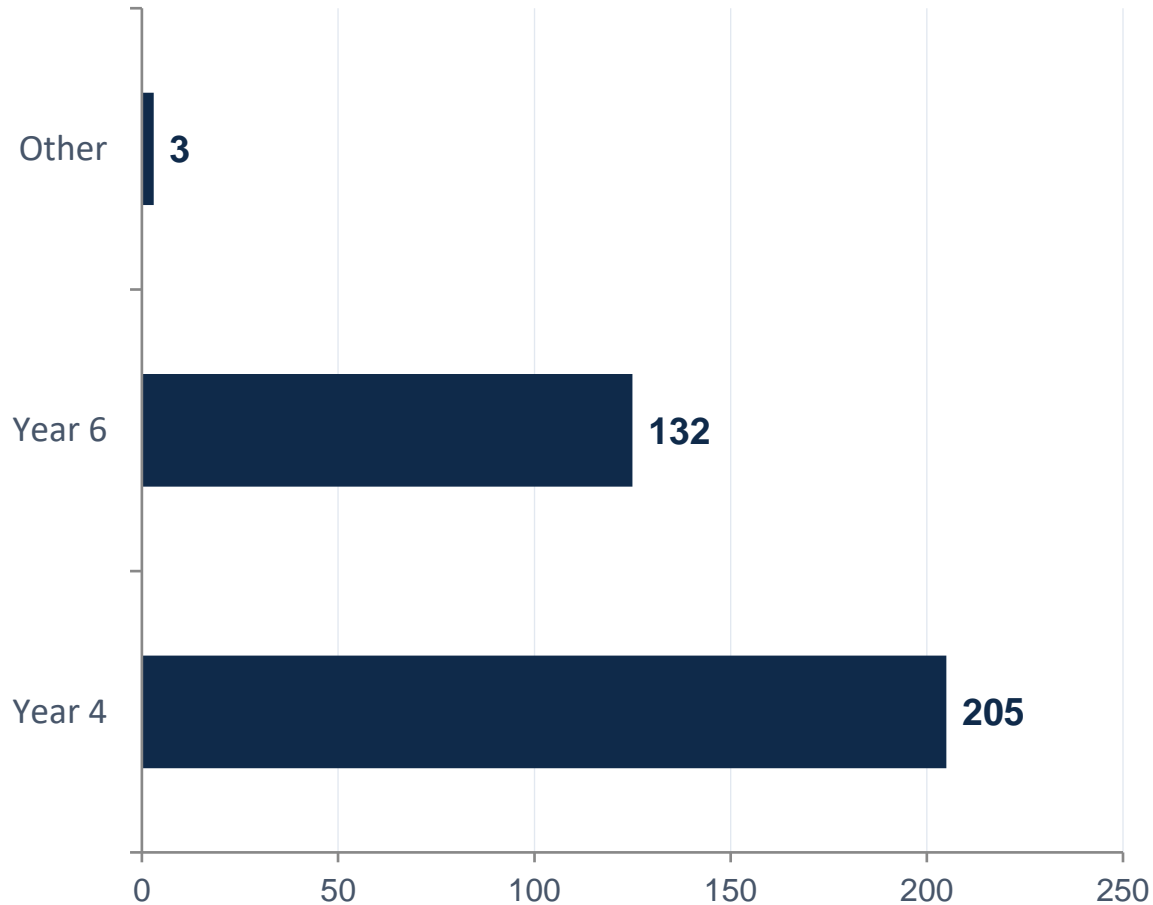
Email  
admin@example.com

Password  
\*\*\*\*\*

Login



# Completion rates: E-modules



## TOTAL RESPONDENTS

**340**

Students (Year 4 + Year 6)

## COMPLETION RATES

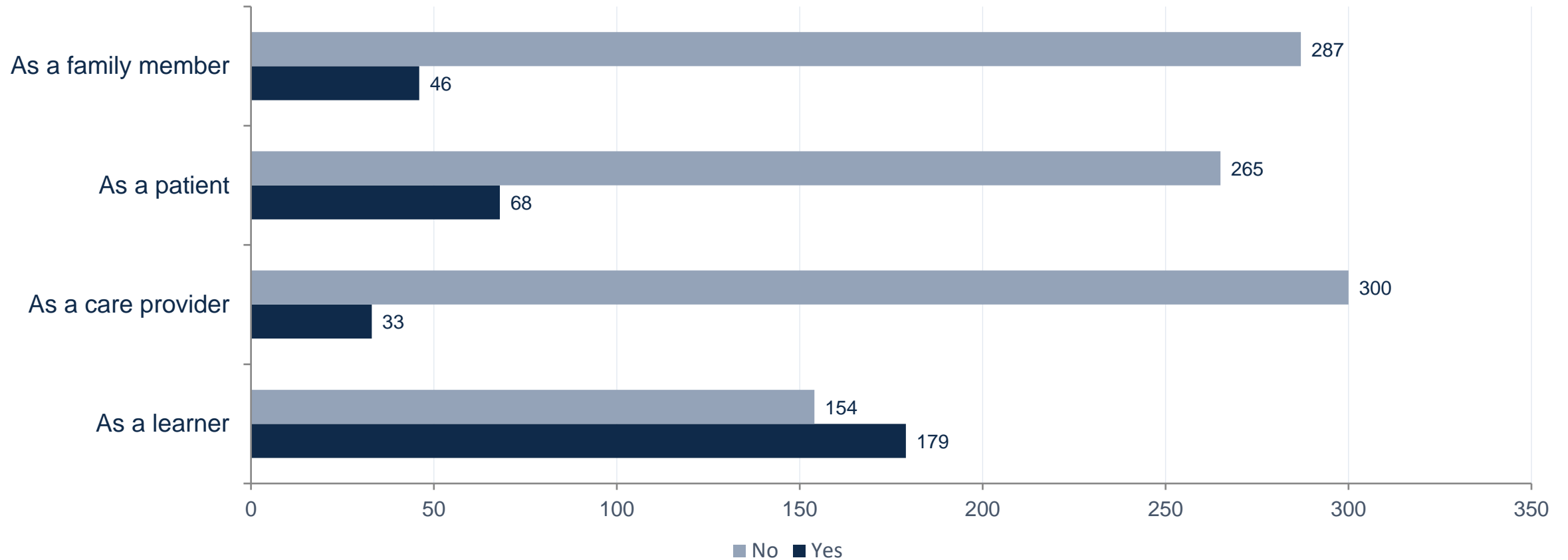
Pre E-Module Survey (340/340)

Post E-Module Survey (320/340)



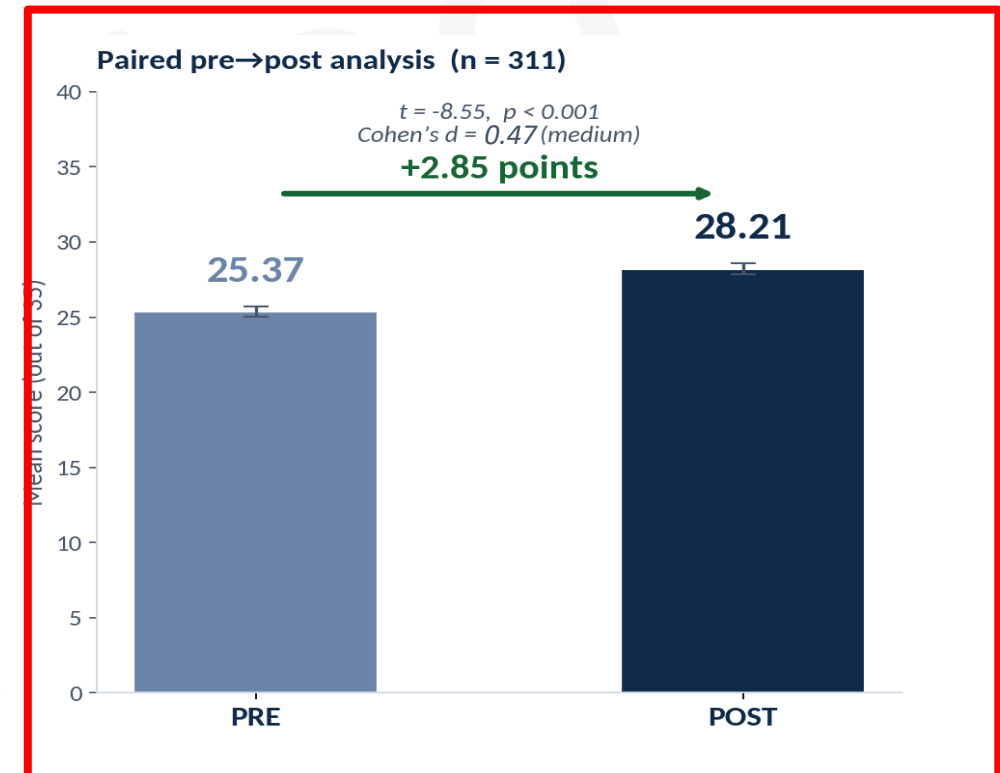
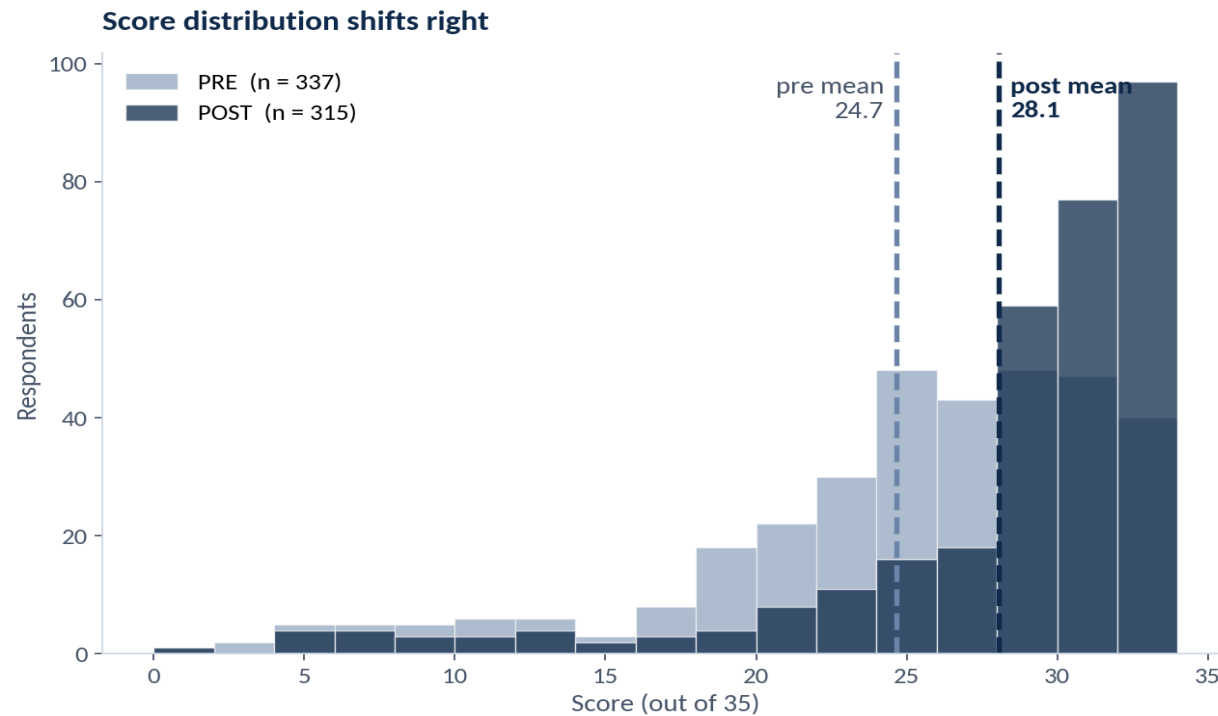
# Prior experience with telemedicine

*Most students have encountered telemedicine as a learner — but **very few have experience as a care-provider, patient, or family member.***



# Students gained telemedicine knowledge: +2.9 points pre→post knowledge quiz ( $p < 0.001$ ), Y4 + Y6

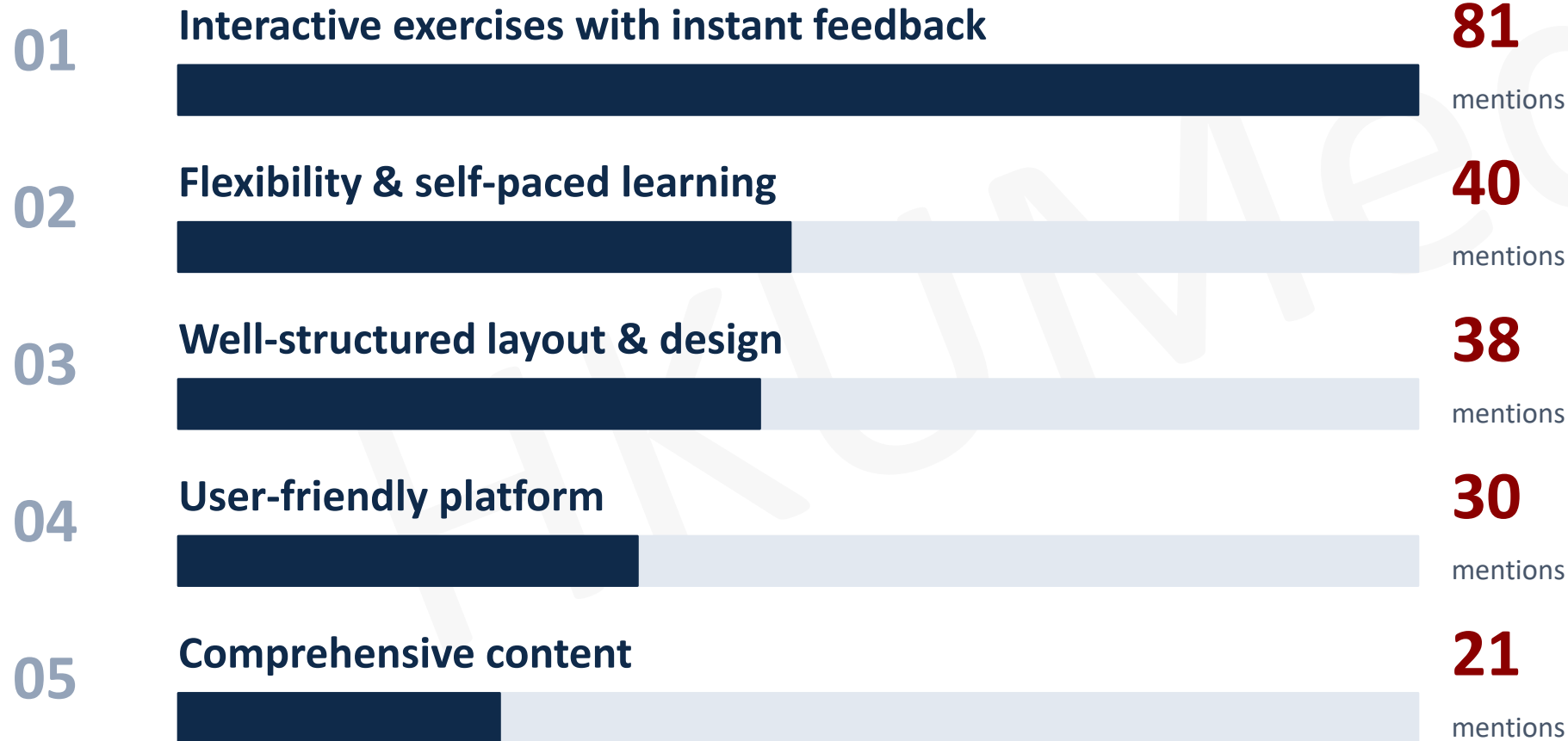
Paired analysis shows a consistent, statistically significant gain with a medium effect size (Cohen's  $d = 0.47$ )  
Gain was slightly larger in Y4 (+3.03) than Y6 (+2.63).



69% of paired students improved · 13% unchanged · 18% declined slightly · paired n = 311



# What students valued most: top 5 themes



## TAKEAWAY

**Interactivity** with feedback dwarfs every other theme — 2x the next strongest signal.

*The focus for any future digital module in FMPC.*



# What students liked least: technical & UX issues

## TECHNICAL · UX

### Difficult navigation

No map; redundant clicks.

### No way to flag questions

Cannot surface confusion.

### AI visuals distracting

Students dislike AI-generated imagery.

### Unclear exercise instructions

Especially drag-and-drop.

### Progress not auto-saved

Window times out; unclear duration.

## ANCHOR QUOTE · S39

*"I do not like AI generated/stock images, they are distracting and giving me a bad impression."*



# What students liked least: pedagogical issues

## PEDAGOGICAL

### No feedback on open-ended answers

No model answer given on purpose.

### Long / repetitive in places

Some content thin, some over-extended.

### No in-person practice

Students want a hybrid component?

### Pre-/post-survey burden

A barrier to learning?

## ANCHOR QUOTE · S21

*"I do not get any opportunities for in-person practice to better prepare for my assessed telemedicine consultation."*



**HKU  
Med**

School of Clinical Medicine  
Department of Family Medicine  
& Primary Care  
香港大學家庭醫學及基層醫療學系

RESULTS

# GenAI feedback

HKUMed



# Technical set up



Home APIs Products Reports Profile Sign out

## Staff GenAI Services (staff-genai-services)

Previously known as the OpenAI Chat (openai-chat) product. Subscribers will be able to access Azure OpenAI, DeepSeek, and Vertex AI Gemini APIs for completions and search.

Staff GenAI Services (staff-genai-services)

▼

```
hku_ai_analysis.py (3.14.4)
from tkinter import messagebox, filedialog
from tkinter.scrolledtext import ScrolledText
from docx import Document
from pathlib import Path
import traceback

CONFIG_FILE = "config.json"

# You can paste your full HKU Chat GPT prompt into DEFAULT_PROMPT below.
DEFAULT_PROMPT = """HKU Chat GPT - gpt5.1
Updated Prompt
Role: You are a clinical communication examiner and experienced family physician
Your task is to analyze a teleconsultation transcript between a medical student
You must provide structured, text-grounded, and rubric-aligned formative feedback
You need to assess then to competency level of a senior medical student (year 4
Give feedback all in English except when quoting directly what students said.

[... paste the entire detailed rubric and formatting instructions here ...]
"""

DEFAULT_API_URL = (
    "https://api.hku.hk/openai/deployments/gpt-5.1/chat/"
    "completions?api-version=2025-01-01-preview"
```

HKU Clinical Communication AI Analysis T...

Student Name:

Paste Zoom Transcript Here:



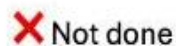
# Example: GenAI feedback on Opening

## Section 1: Opening (Score: 3/5)



Done

- **Introduced self clearly as a medical student**
  - 「咁早晨啦先生，咁我就姓陳嘅，係香港大學醫科學生嘅。」  
→ Clearly identifies as HKU medical student.
- **Explained duration and that it would be recorded; obtained consent to record**
  - 「今日就會進行一個十分鐘嘅問證啦，咁係會有錄影嘅，請問你同唔同意呀？」
  - Patient: 「額。同意。」  
→ Explicit recording consent.
- **Explained limitations of telemedicine (privacy risk & limits of examination)**
  - 「今次係網上...用嘅係加密嘅平臺...冇你同意係唔會錄影或者截圖嘅，咁但係都有少少風險就係啲資料外泄啦。」
  - 「我哋今次用視像呢就做唔到一啲身體上面嘅檢查嘅，咁如果有需要嘅話呢，就要需安排呢個面診嘅...」  
→ Covers privacy risk and inability to perform physical exam; mentions need for inperson visit if necessary.
- **Explained how to reconnect if the call drops**
  - 「如果有電話中斷嘅話呢，我會再打俾你啦。」
- **Obtained consent for telemedicine**
  - 「如果你同意嘅話呢就會繼續我哋呢個錄像咁樣問診嘍。」
  - Patient: 「哦。同意。」



Not done

- **Asked if the patient can see and hear clearly**
  - No question such as 「你聽唔聽到？」 / 「睇唔睇到？」.
- **Assisted with adjusting lighting, sound, or device settings**
  - Not assessable from transcript — no relevant dialogue found.
- **Asked for the patient's name / preferred form of address**
  - Addressed as 「先生」, but no explicit question like 「點稱呼？」.
- **Asked for patient's home telephone number**
  - Not done.
- **Asked for patient's address / confirmed current physical location in Hong Kong**
  - Neither address nor “你而家喺香港邊度？” was asked.
- **Request to see the patient's HKID to visually confirm identity**

# Example: GenAI feedback on History-Taking

## Section 2: History Taking (Score: 3/5)

(Recognising this is a “resultgiving / new diagnosis of diabetes” consult, not a fresh symptombased presentation.)

✔ Done

- **Open question to start the interview / explore reason for visit**
  - 「陳先生知唔知道今次額。睇醫生係咩事呀？」
  - Patient: 「...叫我過嚟睇報告啦，唔知咩事㗎喎。」  
→ Appropriately explores patient’s understanding of why he is here.
- **Clarified what tests were done and what they showed (lab “HPI”)**
  - 「包括咗空腹嘅血糖...血脂肪...腎功能...尿嘅。」
  - 「有兩樣嘢呢...比較擔心嘅嘅就係你個空腹血糖...一次呢就係七點四...第二次呢就係七點五...過咗七呢...屬於係糖尿病㗎啦。」  
→ Clear explanation of tests and results.
- **Explored patient’s immediate reaction/idea about the result**
  - Patient: 「你個零點四咁少，可唔可以咁少？唔當住得唔得呢？...差少少咋喎。」

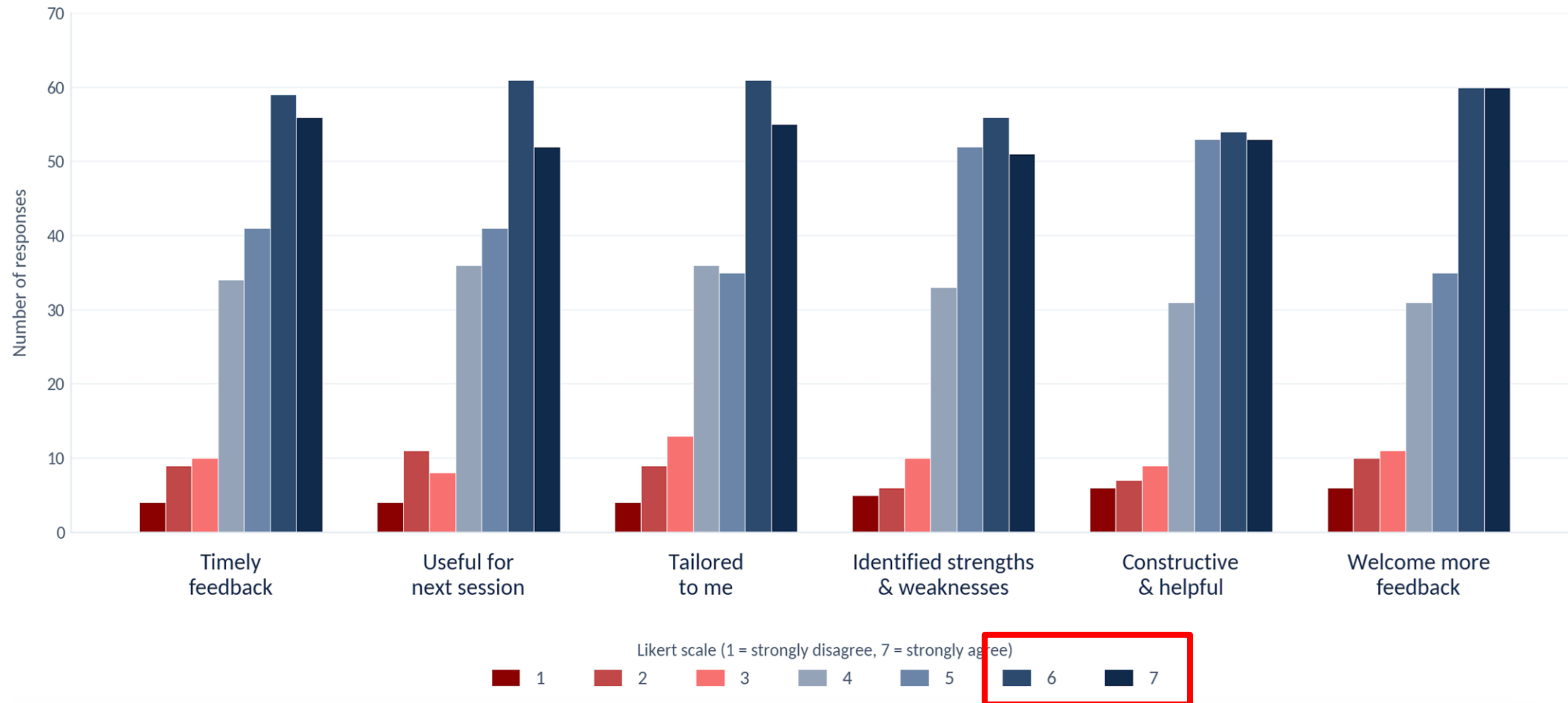
✗ Not done / Improvement needed

- **Symptom review for diabetes (HPI not taken)**
  - No questions on classic DM symptoms:
    - Polyuria (小便多咗？夜尿？)
    - Polydipsia (特別口渴？)
    - Polyphagia (特別肚餓？)
    - Weight change (體重上升或下降？)
    - Visual disturbance, recurrent infections, fatigue, neuropathic symptoms.
  - For a new diagnosis, even when detected by screening labs, a brief symptom review is important.
- **Onset, progression, severity – for hyperglycaemia symptoms**
  - Not asked, because no symptom HPI was taken.
  - Numeric severity (0–10) is **not really applicable** here (lab result discussion rather than pain/dyspnoea).
- **Past medical history (PMH)**
  - No question such as 「平時有冇其他長期病，例如高血壓、心臟病、腎病？」
  - Given new diabetes and positive family history, this is a significant omission.
- **Medication list (names, including for other conditions)**
  - No question like 「有冇長期服藥？食緊啲咩藥？」



# Student rating of GenAI feedback (Year 4+6 combined=213)

7-point Likert scale. Distribution skews strongly toward 5–7 across all items.





# GenAI feedback: strong endorsement, low resistance

END-OF-CURRICULUM SURVEY · Q5\_8 ENDORSEMENT · n = 207

**70%**

find GenAI feedback useful

*("agree", "somewhat agree", or "strongly agree")*

*only*  
**14%**

would actively remove it

*(disagree it is useful)*

Y4: 73% agree useful · Y6: 62% agree useful · Combined: 70% agree, 16% neutral, 14% disagree

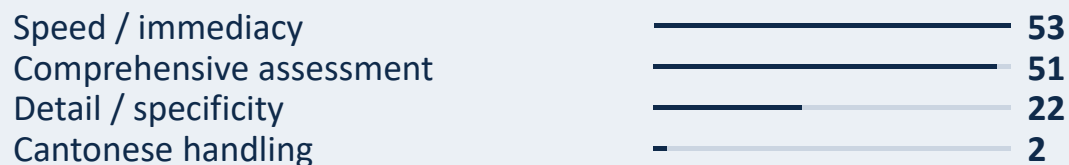


# What students said about GenAI feedback

Top themes and a featured quote from each open-text item (n = 367 responses, Y4 + Y6 combined).

## WHAT WORKED

### Q10 · Best thing about GenAI feedback



- “ It transcribed what I said perfectly, even the cadence and rhythm. — Y4 student
- “ Immediately after 1st one so I can improve in 2nd one — Y4 student
- “ Comprehensive, feedback on every sentence — Y6 student

## WHAT DIDN'T

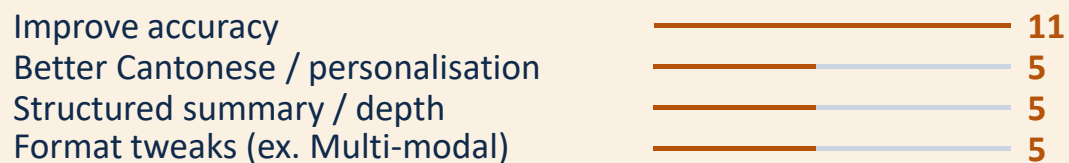
### Q11 · Worst thing about GenAI feedback



- “ It cannot assess the actual tone of the student doing the teleconsultation. — Y4 student
- “ Didn't pick up on all the parts and might misclassify — Y6 student
- “ Useful but still generic AI feedback — Y4 student

## WHAT TO CHANGE

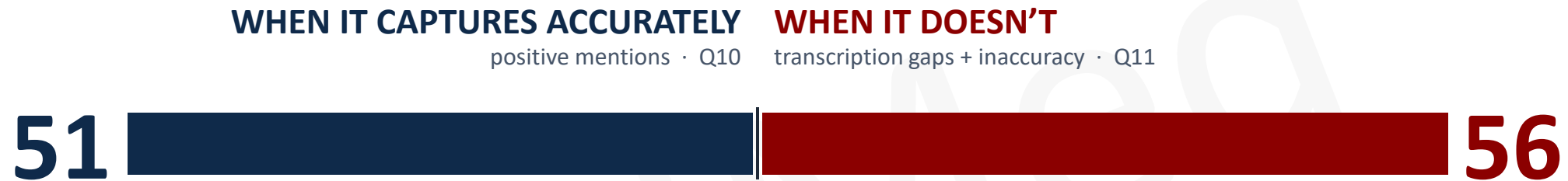
### Q12 · One thing I would change



- “ provide us with the framework of how the AI is comparing us to so we have something to refer to and learn from — Y4 student
- “ More personalised, and perhaps more tailored to Cantonese — Y6 student
- “ Make a summary list at the end for things I did well and things I need to improve — Y4 student

# Accuracy: the strongest positive AND the strongest negative

*The same property students praise most is the one they complain about most.*



“ *It transcribed what I said perfectly, even the cadence and rhythm.*  
— Y4 student

“ *It didn't detect some of the words I said.*  
— Y4 student

What does this mean?

- Transcription accuracy is a barrier to acceptance. When GenAI captures well, it works; when it doesn't, students notice it.
- Should improve over time as technology / infrastructure evolves.



**HKU  
Med**

School of Clinical Medicine  
Department of Family Medicine  
& Primary Care  
香港大學家庭醫學及基層醫療學系

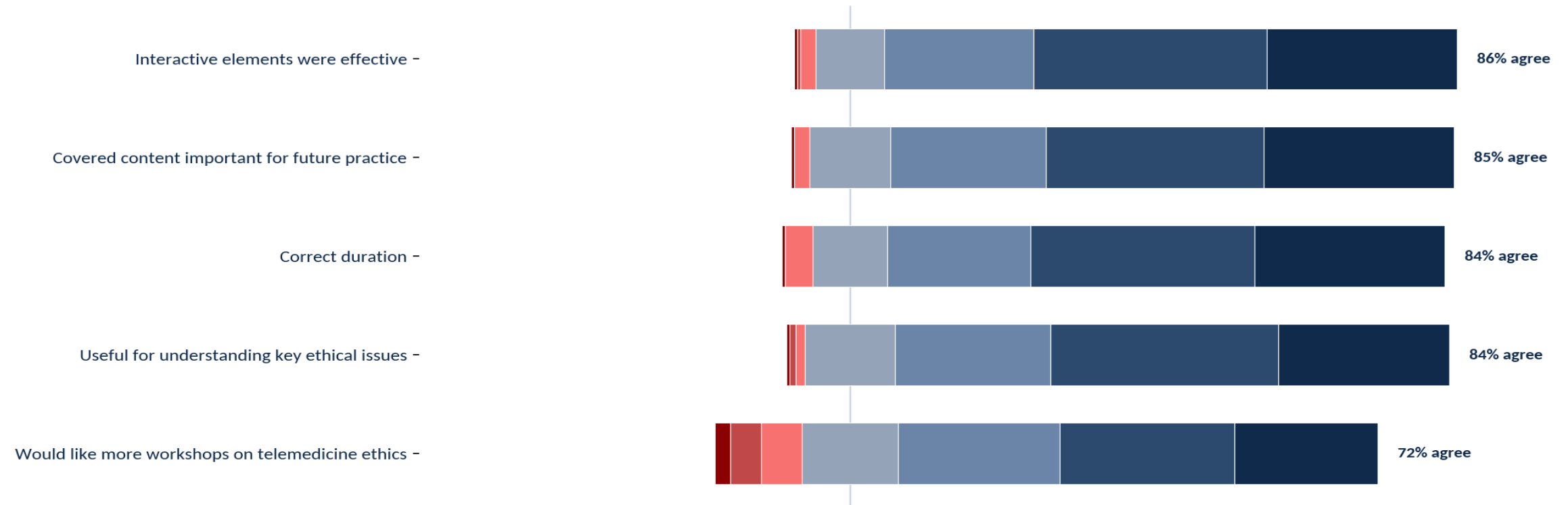
RESULTS

# Ethics workshop & Overall Curriculum



# Ethics workshop: strongly endorsed across both cohorts

Five 7-point Likert items. “Agree” combines somewhat-agree, agree, and strongly-agree (scale points 5–7). The lowest-rated item (72%) still shows that most students want MORE ethics teaching.



Y4 n=148 · Y6 n=65 · combined n=213

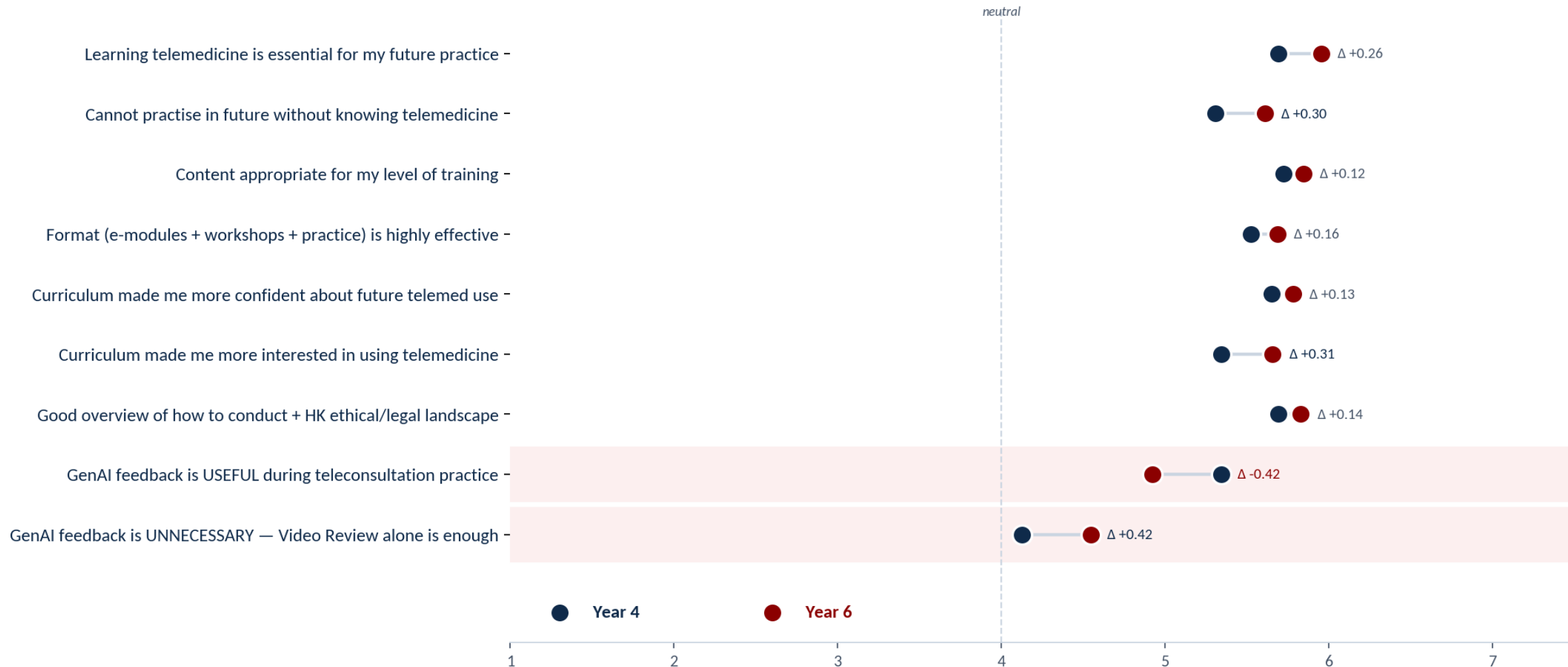




# Curriculum overall: Y6 slightly more enthusiastic — except on GenAI feedback

Mean score per item, 7-point scale. Y4 and Y6 align closely on 7 of 9 items; the two GenAI-feedback items diverge (Y6 less convinced).

END-OF-CURRICULUM SURVEY · Y4 n=143 · Y6 n=64

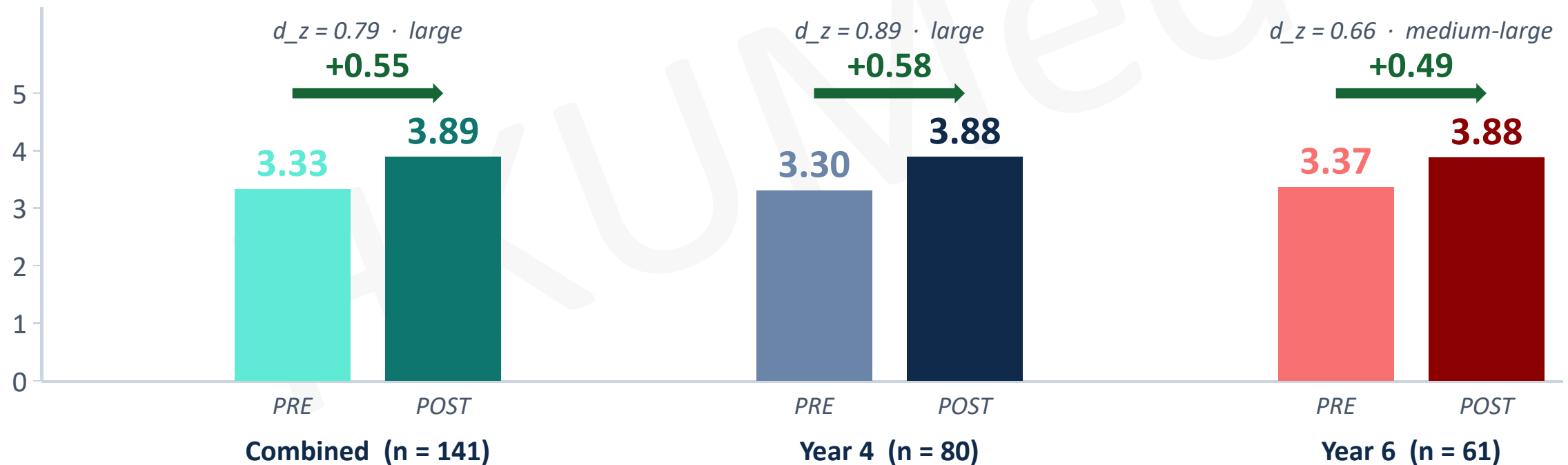




# Self-efficacy/Confidence about teleconsultations gained more than knowledge: +0.55 on 5-pt scale (Cohen's $d = 0.79$ )

10-item GSES-T (General Self-Efficacy Scale, adapted for telemedicine; Schwarzer & Jerusalem). Paired analysis. Composite score = mean of 10 items.

PRE / POST SELF-EFFICACY · GSES-T · Y4 + Y6



73% of paired students improved · 16% unchanged · 11% declined ·  $p < 0.001$  for every group



# Overall Curriculum: what worked, what didn't, what to change

Coded from 207 substantive free-text responses across six prompts ("best / worst / one thing I'd change" for the ethics workshop and the whole curriculum). Approximate theme counts shown.

## TOP THEMES • "Best thing"

### WHAT STUDENTS VALUED

- **Hands-on telemedicine practice**  
*#1 positive in both cohorts* ≈41
- **GenAI feedback (named)**  
*"love the AI feedback"* ≈17
- **Case-based ethics workshop**  
*Kahoot, Menti, real cases* ≈31
- **Variety of teaching modalities**  
*mixed-modality combo* ≈12

## TOP THEMES • "Worst thing"

### WHAT DIDN'T WORK

- **Video Review tutor inconsistency**  
*duplicates AI feedback* ≈7
- **Sharing recordings with peers**  
*"embarrassing"* ≈5
- **Unclear pre-session briefing**  
*"thrown into the ocean"* ≈8
- **e-module ↔ workshop overlap**  
*Y4 + Y6 both noted* ≈5
- **Operational issues**  
*cold air-con, Zoom, schedule* ≈11

## TOP THEMES • "One thing I'd change"

### TOP REQUESTED CHANGES

- **More practice / longer sessions**  
*Y4: 6 min too short* ≈9
- **Fewer cases / shorter workshop**  
*Y6: 12 cases too many* ≈7
- **Don't share videos with peers**  
*keep recordings private* ≈5
- **Clearer pre-session briefing**  
*instructions + expectations* ≈6
- **Better Video Review timing**  
*between sessions* ≈4

## Five lessons from the telemedicine pilot

**1**

**Knowledge and Self-Efficacy/Confidence improved, but confidence improved more**

+0.55 on 5-pt scale ( $d = 0.79$ , large)  
vs +2.9 knowledge points ( $d = 0.47$ )

**2**

**Interactivity with feedback beats every other design choice**

If you optimise one element of a digital module, make it interactive.

**3**

**Students accepted AI-generated feedback with tutor over-sight**

70% of students endorsed GenAI feedback; only 14% would remove it. Accuracy was the largest moderate of acceptance.

**4**

**Your strongest feature can be your weakest feature**

Accuracy of GenAI praised 51× and criticised 56×. When it captures well, students love it; when it doesn't, they notice instantly. Transcription reliability is a rate-limiting step for student trust (potentially addressed with funding / evolving technology)

**5**

**Mind operational friction points**

Y6 uptake lagged Y4 —the two cohorts differed in workflow (Y4 on-site, Y6 at-home with manual upload) as well as in stage of training. Operational friction may lead to reduced uptake, though cohort-stage effects cannot be ruled out.

# References



1. Canadian Medical Association. Report of the Virtual Care Task Force.
2. AAMC. Telehealth Competencies Across the Learning Continuum.
3. Canadian Medical Association. Virtual Care in Canada: Progress and Potential.
4. Royal College of Physicians and Surgeons of Canada. CanMEDS Framework.
5. Recent telemedicine education review (Med Teach, 2024).
6. HKSAR Government. Primary Healthcare Blueprint, 2022.