AI Where It Counts

Human-AI Collaboration in Clinical Image Translation

Jianing Wang | Sr. Scientist II, Statistical Programming, BARDS | MSD, China

PHUSE SDE 2026 · Beijing

MSD

Concrete Evidence: WeChat Translation Failures

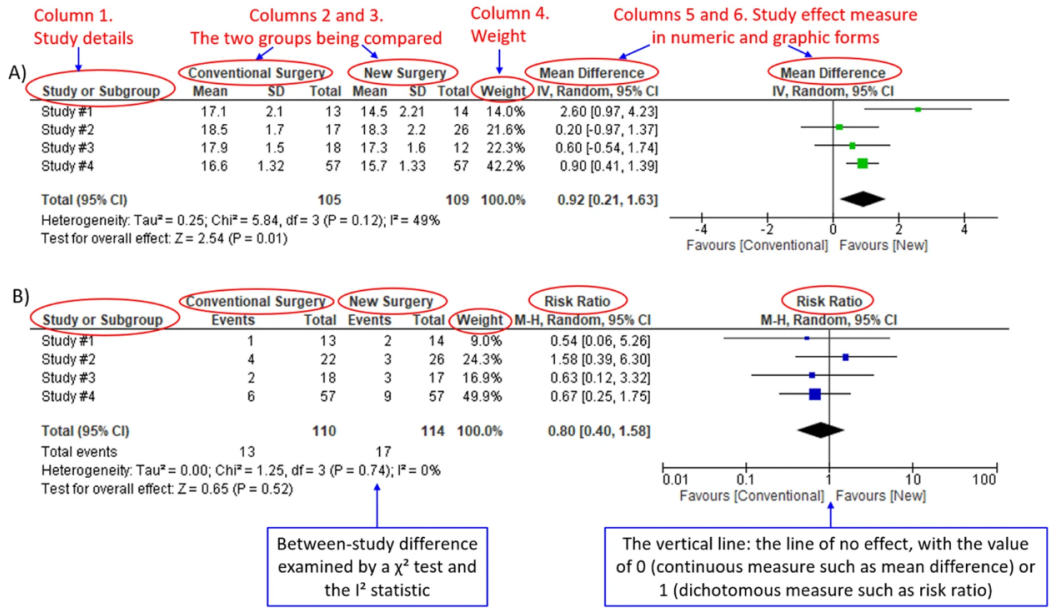
❌ Wrong Terminology: 'Mean' → '吝啬的'

❌ Superscripts destroyed: Tau², Chi², I²

❌ Font chaos + Graphical damage

Fig. 1: Anatomy of a forest plot.

From: The 5 min meta-analysis: understanding how to read and interpret a forest plot

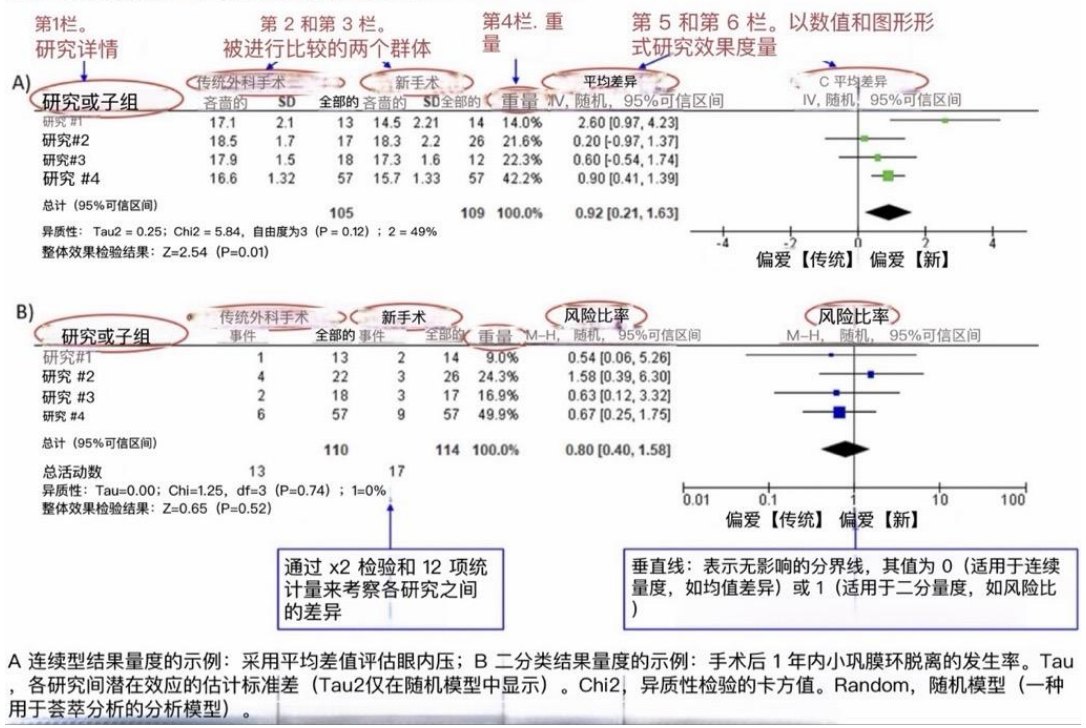


A Example of a continuous outcome measure: Intraocular pressure assessed with mean difference; **B** Example of a dichotomous outcome measure: Incidence of minor zonulolysis, at 1 year after surgery. Tau, the estimated standard deviation of underlying effects across studies (Tau^2 is only displayed in the random model). Chi^2, the value of Chi-square test for heterogeneity. Random, random model (an analysis model in meta-analysis).

Original

图 1：森林样方解剖图。

来源：5分钟元分析：学会如何解读和解释森林图



A 连续型结果量度的示例：采用平均差值评估眼内压；B 二分类结果量度的示例：手术后 1 年内小巩膜环脱离的发生率。Tau，各研究间潜在效应的估计标准差（Tau2仅在随机模型中显示）。Chi2，异质性检验的卡方值。Random，随机模型（一种用于荟萃分析的分析模型）。

WeChat Translation

Source: Chang Y, et al. Eye 36, 673-675 (2022)

Clinical Images: What Makes Them Special?

- Lessons from WeChat failure → requirements for a proper solution

Favorable Characteristics

- Text is typeset, not handwritten→ high OCR accuracy
- High-resolution source documents
- Minimal distortion or screenshot artifacts
- Structured layout (axes, legends, footnotes)

Hard Requirements

- Terminology precision and consistency
- Editable output — human can review & edit
- Reproducible — same input → same output
- Auditable — translation provenance must be traceable

MSD

Two Roads: Inpainting vs Canvas Overlay

Commercial Inpainting

Detect → Erase pixels → AI regenerates → Flatten

Destructive | Irreversible | Static

Original pixels are gone. No undo.

vs

Canvas Overlay

Detect → Keep original → Overlay text → Editable

Non-destructive | Reversible | Editable | GxP-traceable

Original image untouched. Always.

Core Design: Keep original image INTACT. Overlay translated text as movable, resizable, correctable blocks.

MSD

Our Approach: The Overall Workflow

Non-destructive canvas overlay with full human editing control

1 Non-Destructive

Original image untouched — always

- **Translated Text** *Editable*
- **Cover Rectangle** *Anchored*
- **Original Image** *Untouched*

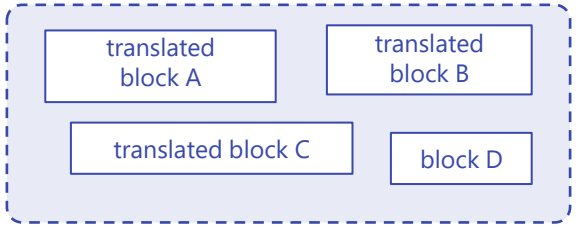
2 AI-Assisted

Dictionary first, AI only when needed

- Dictionary Exact Match — no AI
- LLM Cache — no AI
- RAG-Enhanced LLM Call — AI

3 Human Authority

Every block fully editable before export



Interactive canvas — edit any block before export

1 Input + Local OCR
Deterministic, 100% local

→

2 AI Translation Draft
1-2 LLM calls per image

→

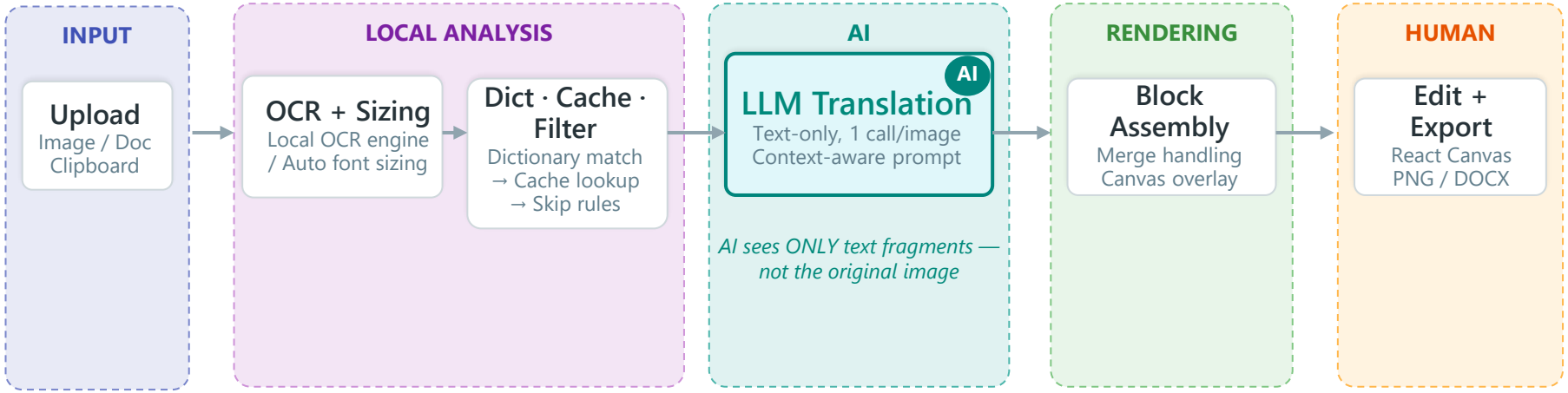
3 Human Edit + Export
Canvas editor → PNG / DOCX

AI generates high-quality draft → Human reviews and edits → Original image always preserved

MSD

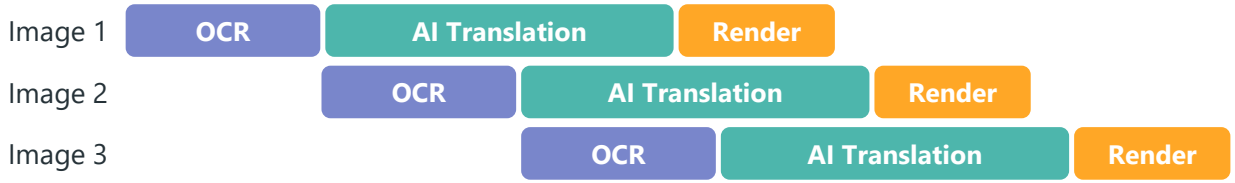
Text-Driven End-to-End Pipeline

1 AI call per image → all other steps are deterministic Python



Pipelined Parallel Processing

OCR sequential → AI parallel across images → Render sequential



Local OCR + Deterministic Font Sizing

PaddleOCR v5 (Local)

- Server models: EN/CH/JP pre-downloaded
- Output per block:
  text + confidence + bbox (x,y,w,h) + angle
- Cached by SHA256(image) :
  Same input → same output, every time
- Zero cloud dependency :Zero network calls

Font Sizing (3 Tiers)

- base_size = median(heights) × correction
- CJK: 0.70 / Latin: 0.80
- Three tiers:
  - ➢ Title: > 1.8× median
  - ➢ Body: default tier
  - ➢ Footnote: < 0.55× median
- Deterministic: same image → same sizes

DATA PRIVACY: Image pixels NEVER leave the server. OCR runs 100% locally. Fully deterministic.

MSD

Block-Aware Translation: The OCR Splitting Problem

OCR sees 3 boxes. A human reader sees 1 sentence.

When Text-Driven was designed, vision models lacked accuracy for text-level detail. Prompt engineering compensated for this limitation.

The Problem

Block 1 | Block 2 | Block 3

Percentage of Subjects | with Treatment-Emergent | Adverse Events

WITHOUT block-aware ❌

OCR sees 3 boxes. A human sees 1 sentence.

“受试者百分比” | “与治疗相关的” | “不良事件”

Each block translated independently — meaning severed

The Solution – AI Prompt with CoT + few shots

Analyze relationships → Combine into sentences → Translate as units → Distribute proportionally → Verify boundaries → Final assembly

AI reconstructs semantic units first, translates them whole, then redistributes back to original block positions.

WITH block-aware ✓

“发生治疗期不良” | “事件的受试者” | “百分比”

Each block translated independently — meaning severed

MSD

Translation Strategy: Dict → Cache → RAG-LLM

Priority Waterfall

1. **Dictionary Exact Match** — PYTHON
   Built-in + User provided Dictionary — AI cannot override

↓

2. **LLM Cache** — PYTHON
   Same text → instant lookup, zero API cost

↓

3. **RAG-Enhanced LLM** — AI
   ChromaDB retrieval feeds terminology into prompt

Smart Filter

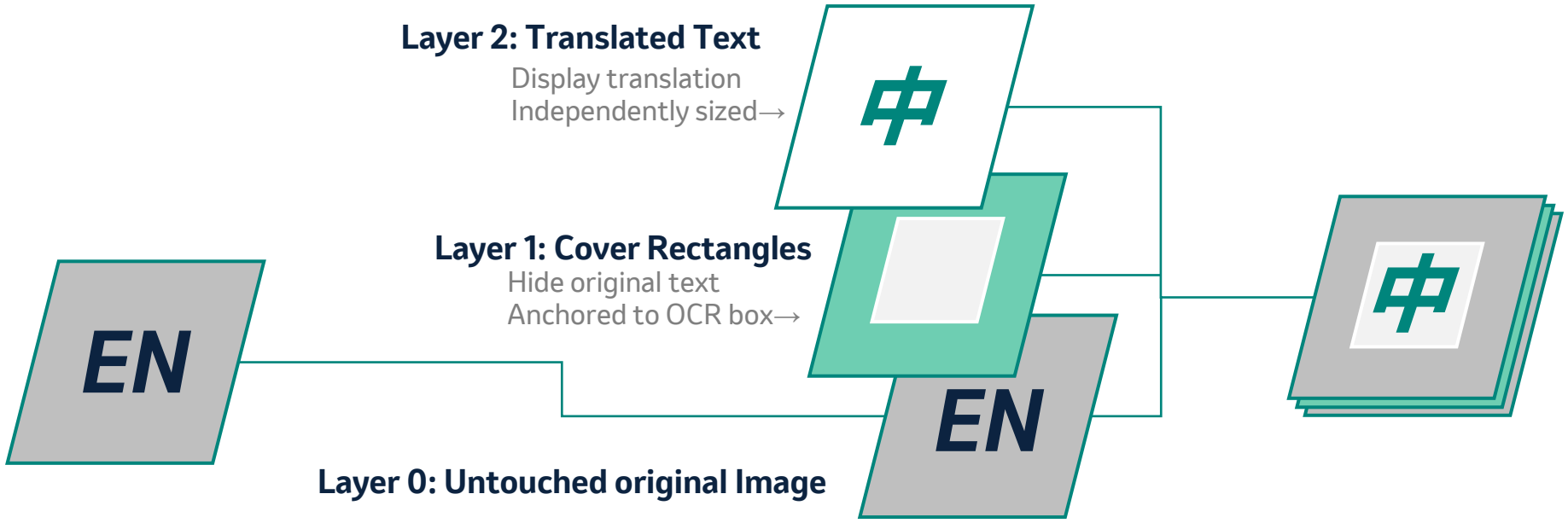
- 74+ regex rules auto-skip non-translatable tokens (N/A, digits, protocol IDs, URLs, symbols …)

Make most blocks resolved without any AI call. Dictionary authority ensures regulatory consistency.

MSD

2-Layer Architecture: Coverage vs Readability

Decouples hiding original text from displaying translation



The Conflict

Cover ≥ original bbox.
But translation may
need different font size.

Single-Layer ❌

↑ font → cover shrinks
→ original shows.
OR force large box
→ wastes space.

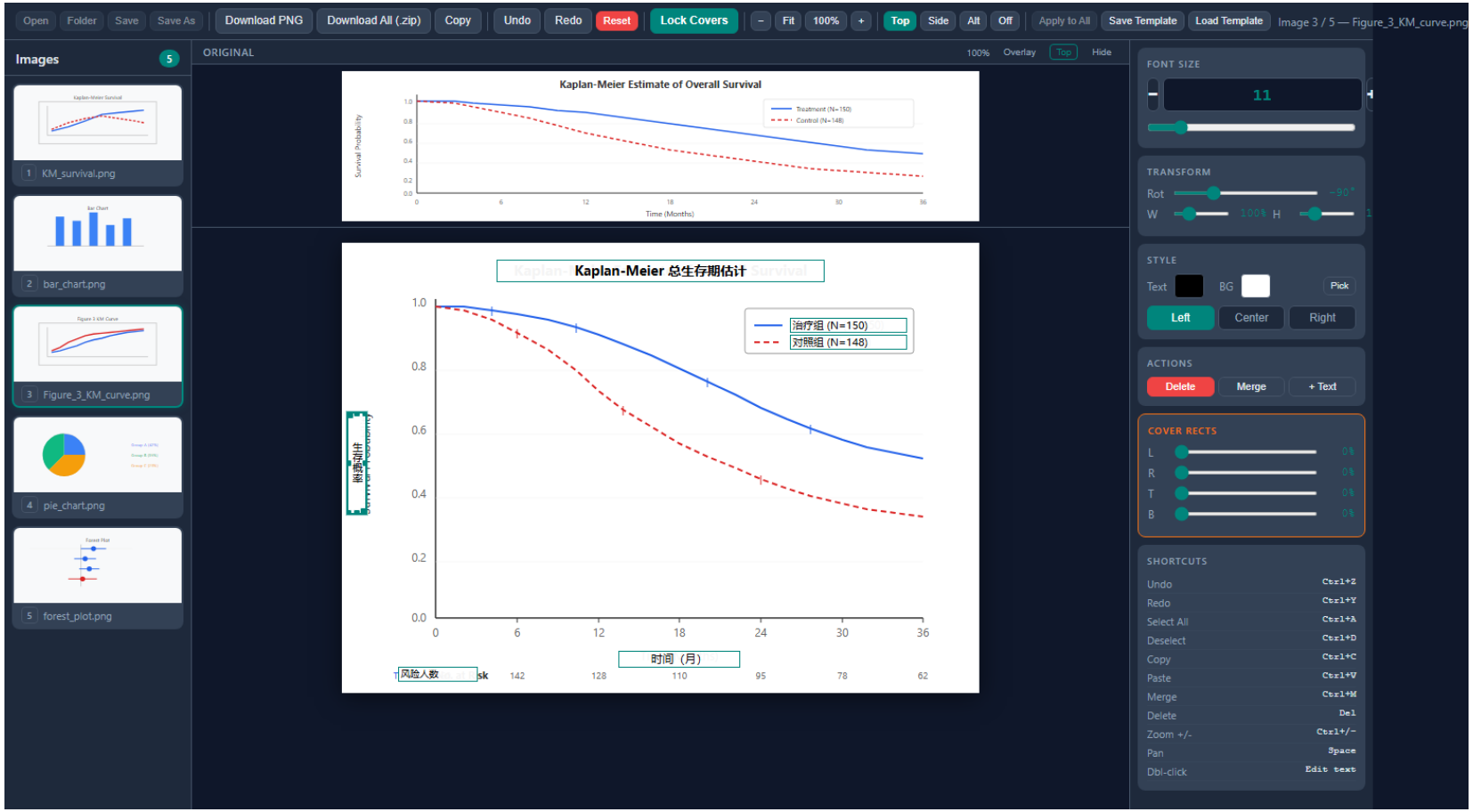
2-Layer Solution ✓

Layer 1: coverage.
Layer 2: readability.
No interference.

MSD

Canvas Editor - The Human Authority Interface

React + Fabric.js → postMessage → Streamlit session_state



Full Editing Control

1. Move, resize, rotate any text block
2. Change font size, color, alignment
3. Merge multiple blocks (paragraphs)
4. Undo / Redo full history
5. Keyboard shortcuts (Ctrl+Z/Y/M/A)
6. Export: PNG, DOCX, Clipboard

All the text blocks are editable

Note: Simulated data for demonstration only

MSD

Export & Traceability: 4-Tier System

PNG

Quick share

Editable?

No

Git-friendly?

Large binary

Compliance

Basic

.gtproj

Full project

Editable?

Yes (re-open)

Git-friendly?

Large binary

Compliance

Full audit trail

.gtdata

Lightweight checkpoint

Editable?

Yes (re-open)

Git-friendly?

Yes (small JSON)

Compliance

Diff-friendly

HTML

Zero-dependency archive

Editable?

Yes (browser)

Git-friendly?

Medium

Compliance

Reviewer distribution

Document Writeback:

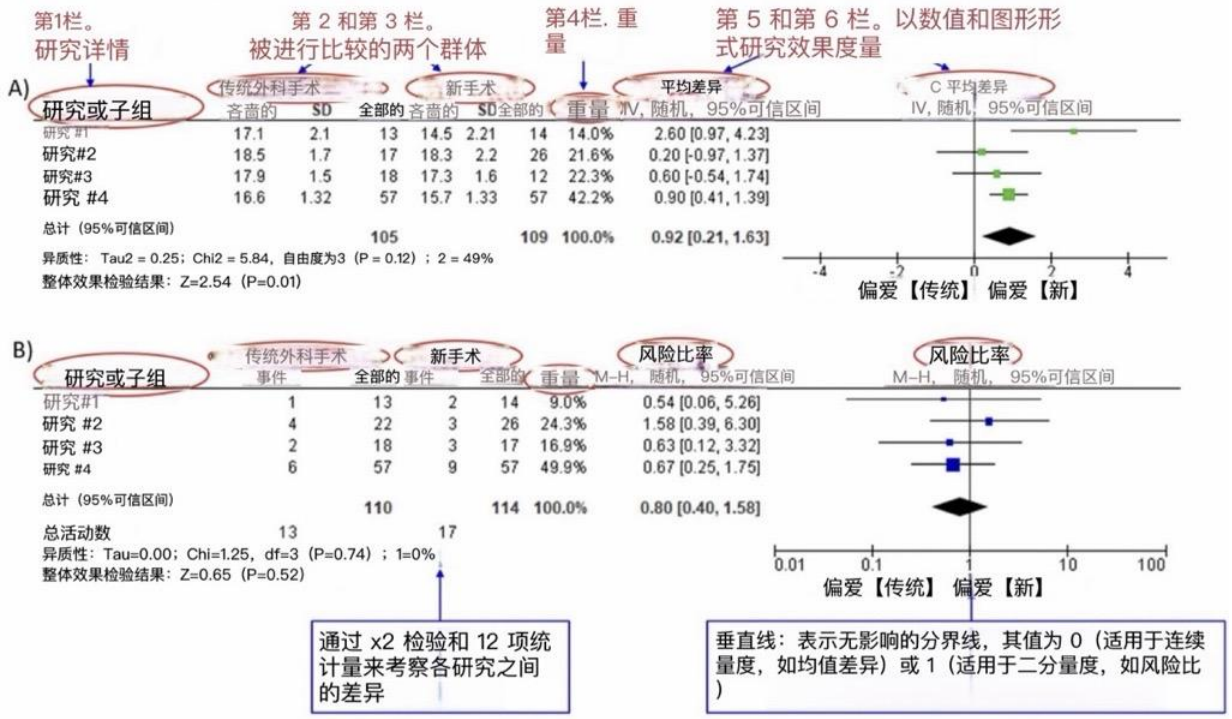
Upload original → Upload translated PNGs → Match by naming → Replace images → Download modified doc

Text-Driven Pipeline Result: Forest Plot Anatomy

WeChat

图 1：森林样方解剖图。

来源：5分钟元分析：学会如何解读和解释森林图

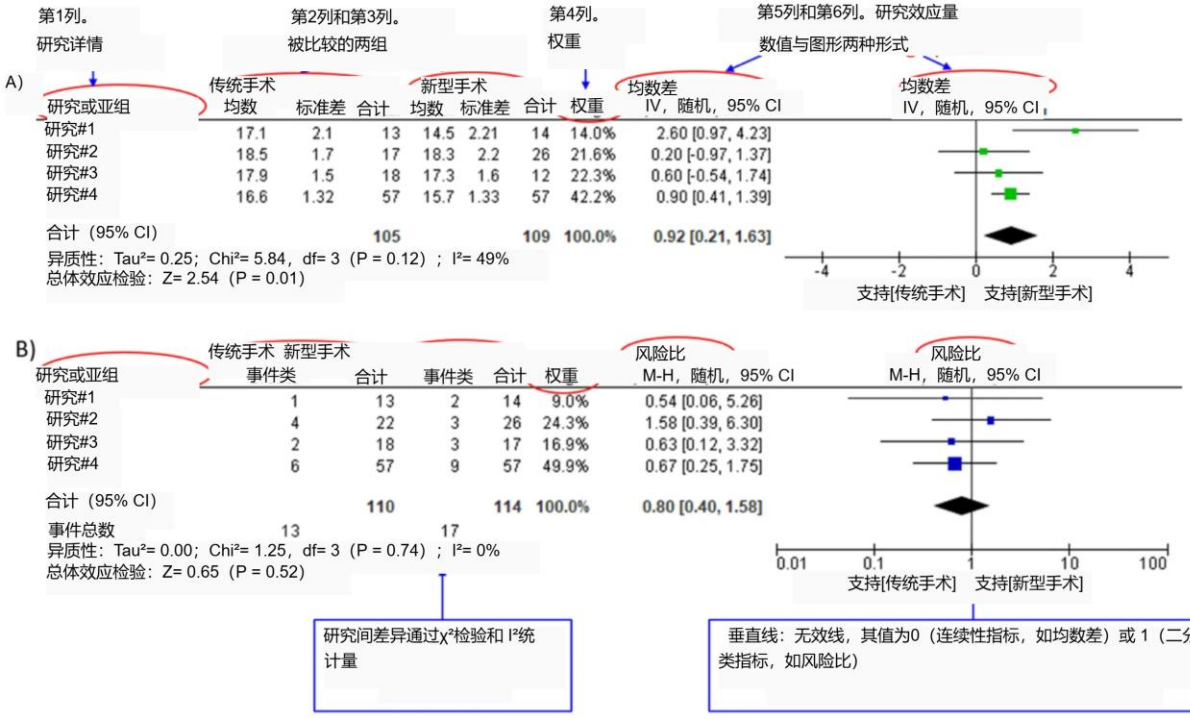


A 连续型结果量度的示例：采用平均差值评估眼内压；B 二分类结果量度的示例：手术后 1 年内小巩膜环脱离的发生率。Tau，各研究间潜在效应的估计标准差（Tau2仅在随机模型中显示）。Chi2，异质性检验的卡方值。Random，随机模型（一种用于荟萃分析的分析模型）。

Text-Driven

图1：森林图的结构解析。

来源：5分钟荟萃分析：如何阅读和解读森林图



A 连续性结局指标示例：以均数差评估眼压；B 二分类结局指标示例：术后1年小带断裂的发生率。 Tau，各研究潜在效应的估计标准差（Tau²仅在随机效应模型中显示）。Chi²，异质性卡方检验值。 Random，随机效应模型（荟萃分析中的一种分析模型）。

MSD

Text-Driven Pipeline Result: Forest Plot Anatomy

WeChat

A)

研究或子组	传统外科手术			新手术		
	吝啬的	SD	全部的	吝啬的	SD	全部
研究 #1	17.1	2.1	13	14.5	2.21	
研究#2	18.5	1.7	17	18.3	2.2	
研究#3	17.9	1.5	18	17.3	1.6	
研究 #4	16.6	1.32	57	15.7	1.33	

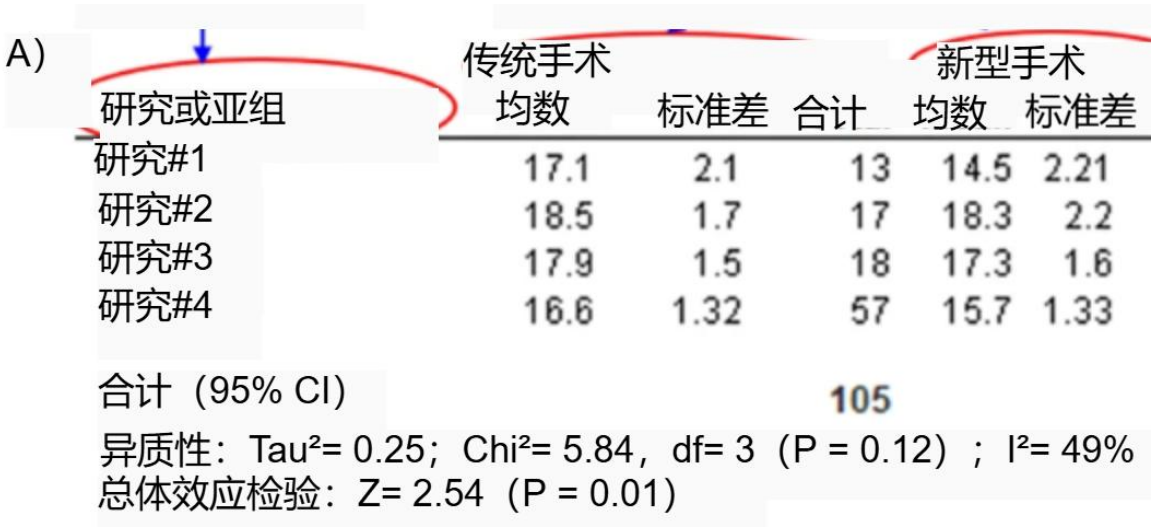
总计（95%可信区间） 105

异质性： Tau2 = 0.25; Chi2 = 5.84，自由度为3（P = 0.12）；2 = 49%

整体效果检验结果：Z=2.54（P=0.01）

- ❌ Wrong Terminology: 'Mean' → '吝啬的'
- ❌ Superscripts destroyed: Tau^2, Chi^2, I^2
- ❌ Font chaos + Graphical damage

Text-Driven



A)

研究或亚组	传统手术			新型手术	
	均数	标准差	合计	均数	标准差
研究#1	17.1	2.1	13	14.5	2.21
研究#2	18.5	1.7	17	18.3	2.2
研究#3	17.9	1.5	18	17.3	1.6
研究#4	16.6	1.32	57	15.7	1.33

合计（95% CI） 105

异质性：Tau²= 0.25；Chi²= 5.84，df= 3（P = 0.12）；I²= 49%

总体效应检验：Z= 2.54（P = 0.01）

MSD

Text-Driven Pipeline Result: PAA Form

Original

PALATABILITY AND ACCEPTANCE ASSESSMENT (PAA)

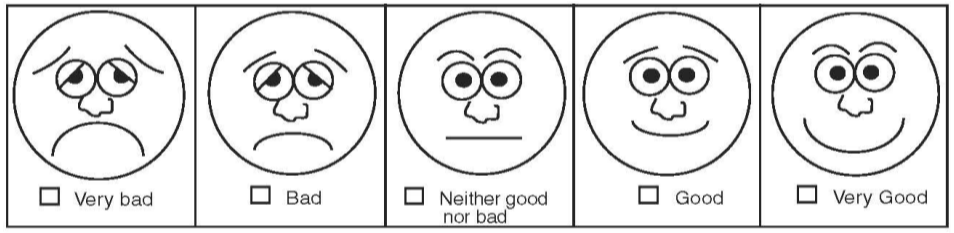
Specify completion date: _____
DD-Mon-YYYY

Please mark one box below to identify the person entering the responses on the questionnaire.

- ☐ The patient
- ☐ The parent/primary caregiver
- ☐ The patient and parent/primary caregiver
- ☐ The health care provider (e.g., physician, nurse, medical assistant or nursing assistant caring for the patient)

This form should be completed at the time of oral medication administration.
Please answer all questions.

1. Please mark the box of one picture below to show how you/the patient feels about the taste of the medication:



2. Please indicate if you/the patient had any of the following problems when taking the medication by mouth:

a. Refusing	☐ Yes	☐ No
b. Spitting Out	☐ Yes	☐ No
c. Throwing Up or Spitting Up	☐ Yes	☐ No
d. Gagging	☐ Yes	☐ No
e. Other, specify	☐ Yes	☐ No

3. Are there any other comments related to the taste of the medication? ☐ Yes ☐ No

Specify comments: _____

I confirm this information is accurate.	Subject's/Caregiver's Initials:	Date:

Text-Driven

适口性和接受度评估（PAA）

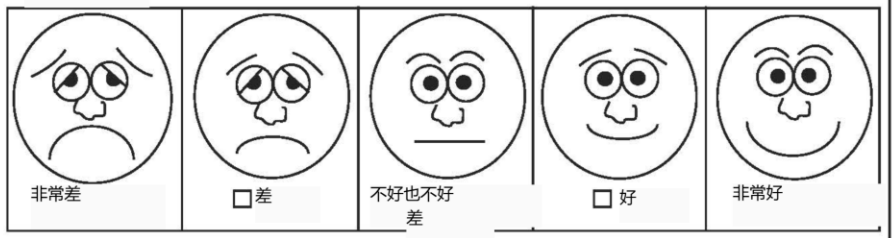
请注明完成日期：
日-月-年

请在下方勾选一个方框，以确定由谁填写本问卷中的回答。

- ☐ 患者
- ☐ 父母/主要照护者
- ☐ 患者和父母/主要照护者
- ☐ The health care provider (e.g., physician, nurse, medical assistant or nursing assistant caring (为患者填写)

本表应在口服给药时填写。
请回答所有问题。

1. 请在下方图片中勾选一个方框，以表示您/患者对该药物味道的感受：



2. 请说明您/患者口服该药物时是否出现以下任何问题：

a. 拒绝服用	是	否
b. 吐出	是	否
c. 呕吐或溢吐	是	否
d. 干呕	是	否
e. 其他，请注明	是	否

3. 是否还有与该药物味道相关的其他意见？

请注明意见：

我确认此信息准确无误。	受试者/照护者姓名首字母：	日期：

MSD

Text-Driven Pipeline Result: Clinical Poster

Original

Results

- As the data cut, a total of 55 patients with advanced cervical cancer had been randomized and received at least one dose of Nilvanstomig 10 or 20 mg/kg. Baseline characteristics are shown in Table 1.
- Among the total 22 patients on the 20 mg/kg group who hadn't treated with any prior ICI, the confirmed ORR was 40.9%(Figure 2), and the DCR was 68.2% per IRC's assessments. The mPFS had not yet been reached per the IRC's assessments(Figure 3), but reached 11.01 months per the investigator's assessments.
- Treatment related adverse events (TRAEs) occurred in 46 (83.6%) patients , with 5 (9.1%) being grade 3-4 (one patient each with hypocalcemia, myositis, rash, hypertension and anemia). Serious adverse events (SAEs) occurred in 8 (14.5%) subjects, with only myositis case assessed as related, which also resulted treatment discontinuation. No grade 5 TRAE encountered. Grade 3-4 immune-related adverse events (irAEs) were observed in 4 (7.3%) patients, including myositis, rash, hypertension and anemia (Table 2).
- Nilvanstomig concentration in serum displayed target mediated drug disposition characteristics at lower doses. At doses of 3 mg/kg and above, C_{max} and AUC_{inf} increased in a near dose-proportional manner (Figure 4). Near saturated receptor occupancy on peripheral T cells maintained at least 3 weeks at 3 mg/kg or higher doses, supporting a dosing interval of 3 weeks or longer (Figure 5). Preliminary immunogenicity results indicated low incidence of anti-ZG005 antibodies.

Figure 2. Best Percentage Change in Target Lesion Size from Baseline within the Cervical Caner Cohort by IRC

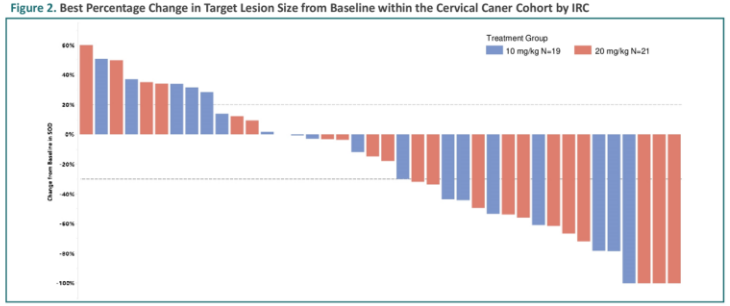
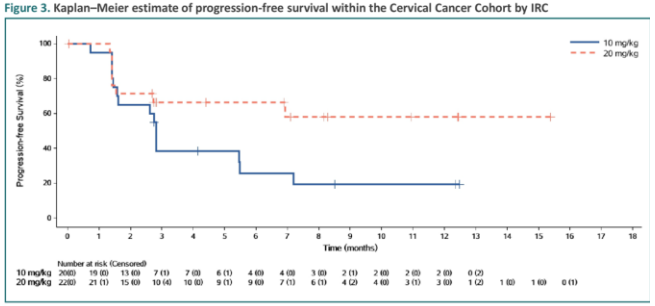


Figure 3. Kaplan–Meier estimate of progression-free survival within the Cervical Cancer Cohort by IRC

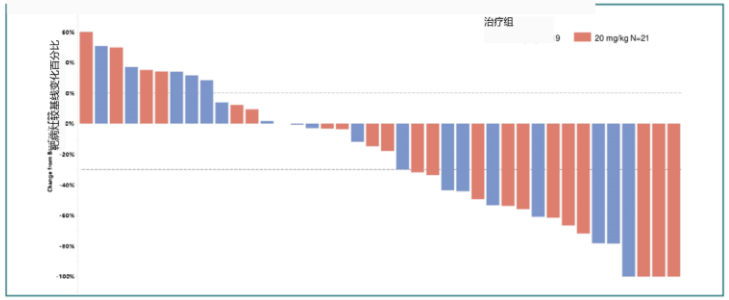


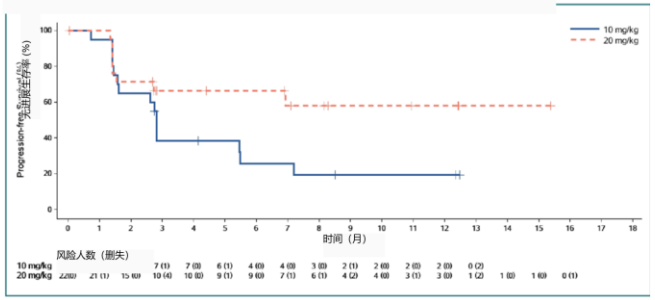
Text-Driven

结果

- 截至数据截止日，共有55例晚期宫颈癌患者已被随机分组并接受了至少一剂Nilvanstomig 10或20 mg/kg治疗。基线特征
  见表1。
- 在20 mg/kg组中未接受过任何既往ICI治疗的22例患者中，经IRC评估确认的ORR为40.9%（图2），DCR为68.2%。
  经IRC评估的mPFS尚未达到（图3），但经研究者评估的mPFS为11.01个月。
- 治疗相关不良事件（TRAEs）发生于46例（83.6%）患者中，其中5例（9.1%）为3-4级（每例分别为低钙血症、肌炎、皮疹、高血压和贫血）。
  严重不良事件（SAEs）发生于8例（14.5%）受试者中，仅肌炎病例被评估为相关，且导致停止治疗。未发生5级TRAE。
  3-4级免疫相关不良事件（irAEs）见于4例（7.3%）患者，包括肌炎、皮疹、高血压和贫血（表2）。
- Nilvanstomig血清浓度在较低剂量时表现出靶点介导的药物处置特征。在3 mg/kg及以上剂量时，Cmax和AUCint呈近似剂量
  比例的方式增加（图4）。外周T细胞上的受体占有率在3 mg/kg或更高剂量下至少维持3周接近饱和，支持3周或
  更长的给药间隔（图5）。初步免疫原性结果显示抗ZG005抗体发生率低。

图2. 经IRC评估的宫颈癌队列中靶病灶较基线的最佳百分比变化

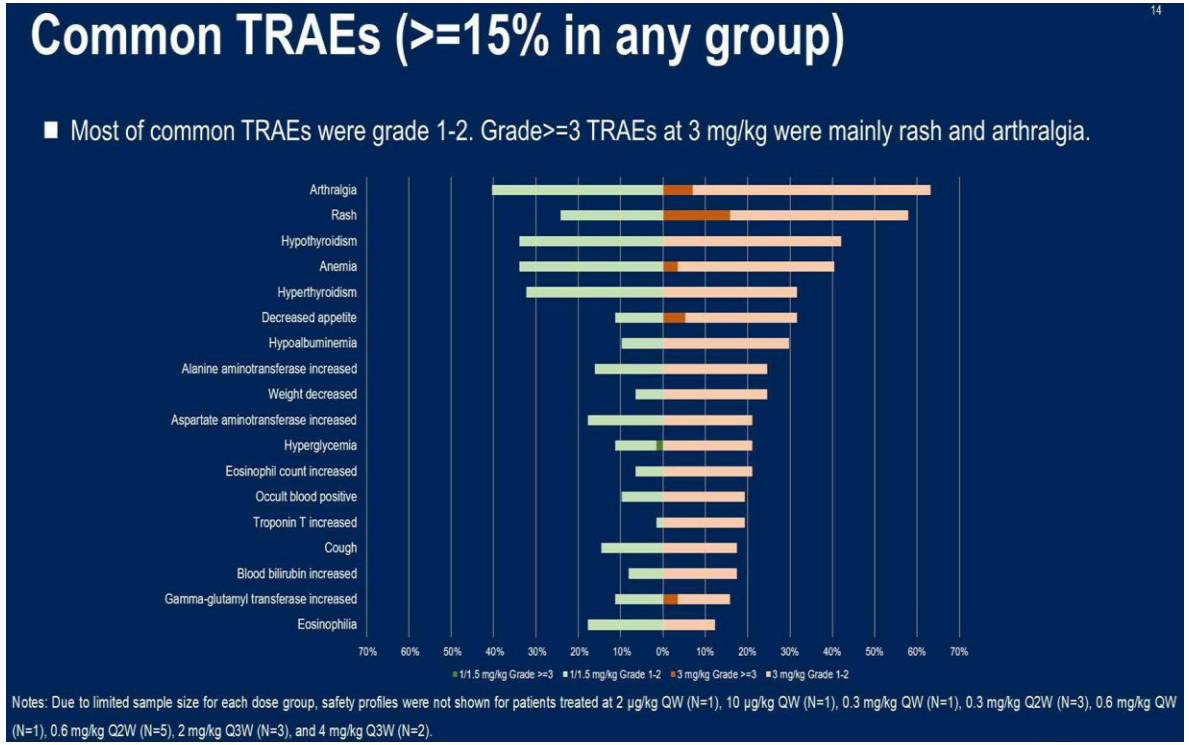




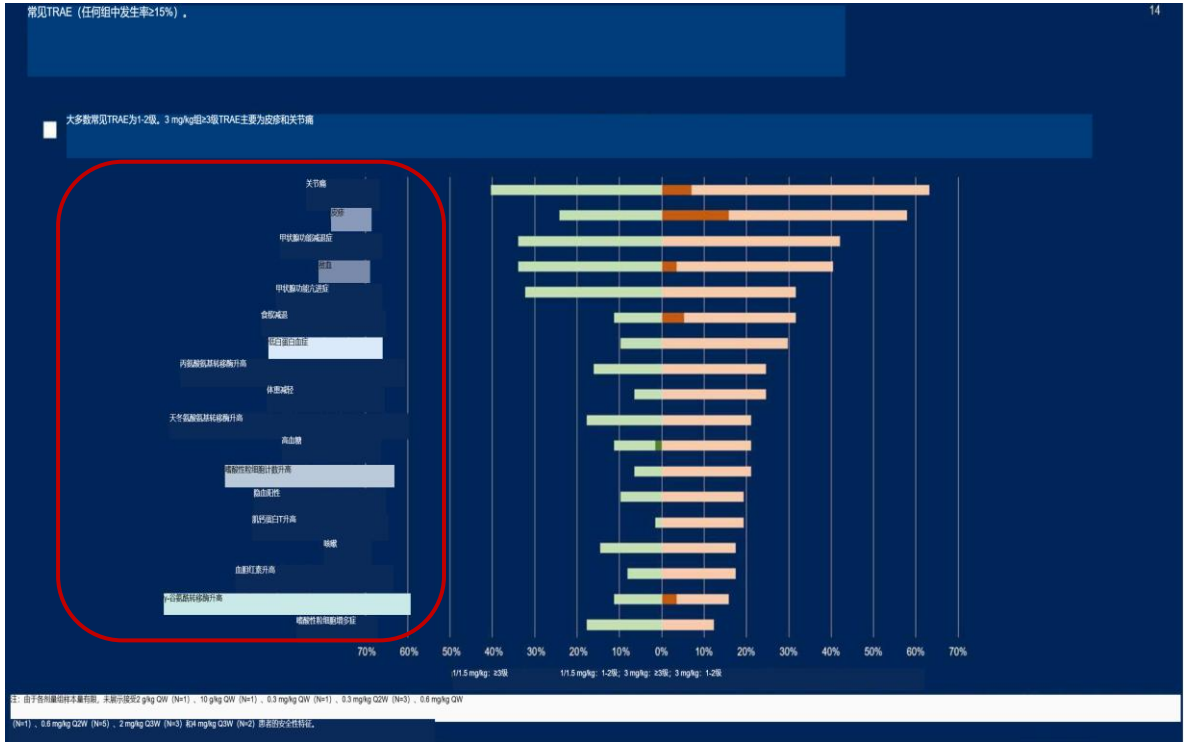
MSD

Text-Driven Pipeline Result: Dark Background Chart

Original



Text-Driven



Text-Driven Pipeline Summary: already Impressive, But…

1	**Fast**	1 AI call per image · speed compounds at batch volume
2	**Data-Sensitive**	AI sees only text fragments— images never leave local server
3	**Deterministic**	Dictionary, 3-tier font, smart Regex filter — all reproducible
4	**Human-Centered**	Canvas editor gives full control over every block

The 2-layer canvas + editor gives users full control.

But a blind model makes avoidable errors —broken alignment, missed context, wrong colors.

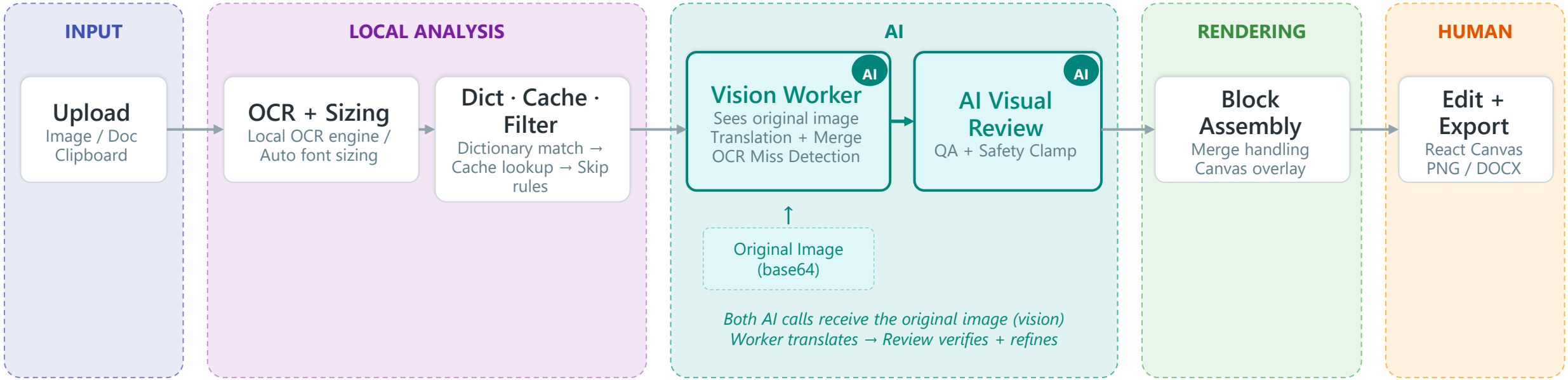
As vision models matured: What if the AI could SEE the original image during translation?

→ Vision-Driven Pipeline

MSD

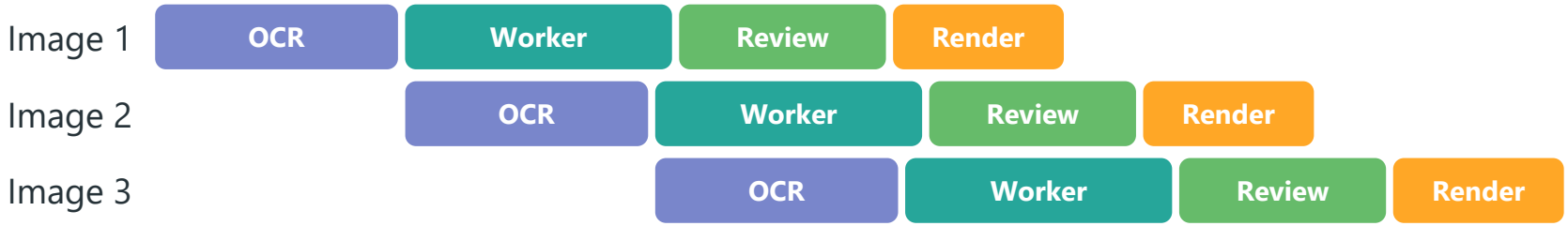
Vision-Driven End-to-End Pipeline

2 AI calls per image — Worker creates, Review verifies. All other steps identical.



Pipelined Parallel Processing

OCR sequential → Worker + Review parallel across images → Render sequential



MSD

Vision-Driven: Core Concept

The **ONLY** architectural change: the translation model now SEES the original image

1 **Original image (base64)**
AI sees the complete figure

2 **OCR blocks in JSON**
Text, coordinates, confidence, font tier

3 **RAG context**
Retrieved terminology from vector database

→

AI Vision
Vision Worker

→

Worker Output

translated_text	Context-aware
merge_group	Paragraph structure detected visually
bg / text_color	Colors determined from image context
missed_blocks	Text OCR failed to detect

Text-Driven

`{text: “Mean", x: 340, y: 82} → translates in isolation`

Vision-Driven

`{text: "Mean", x: 340, y: 82} + image → understands visual context`

MSD

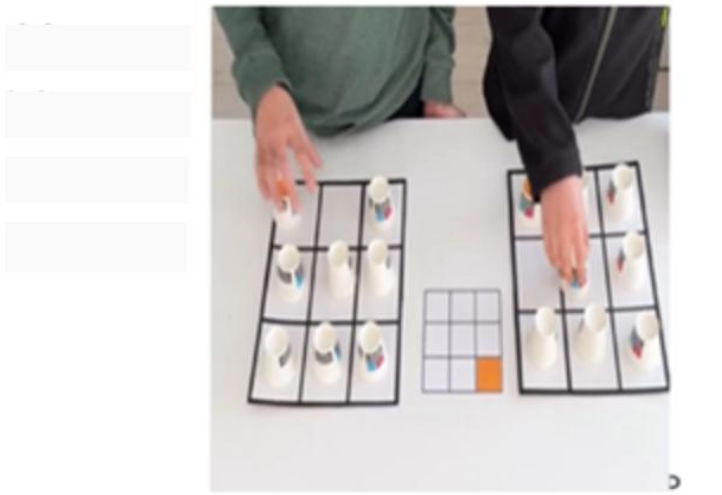
OCR Miss Recovery (Vision-Driven Exclusive)

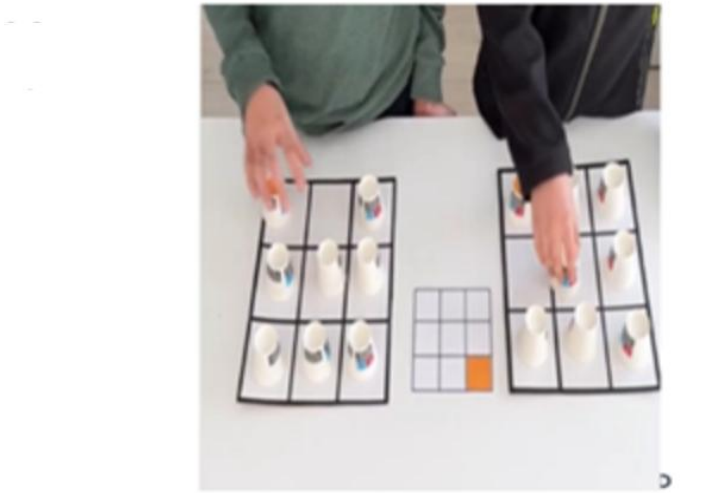
Text-Driven pipeline has zero ability to recover from OCR misses

Low contrast / rotated text → PaddleOCR misses → Vision Worker spots → Python Rule Validation → Inject misses

Python Rule Validation:

① High AI self-confidence | ②Coordinates within image | ③Valid box size | ④Reasonable shape | ⑤Not already found by OCR | ⑥Worth translating | ⑦Hard cap: max injected block





MSD

AI Visual Review: Catching Mechanical Errors Automatically

Worker's remaining errors are predictable — Review fixes them within Python bounds

AI Visual Review

3 Inputs:

- Original image (base64)
- Pillow preview with translations
- Structured block data (JSON)

Outputs:

Per-block:
font size/ width/ height adjustments
+ text_correction + merge_groups

Hard Clamp

Font Size	±30%
Position	±15px
Dimensions	±20%

AI SUGGESTS, Code ENFORCES

Before & After:

Font Uniformity

Before: Inconsistent font size (14px, 18px, 14px)

After: Unified within ±30% (15px, 15px, 15px)

Color Correction

Sample Text — Before: **Unreadable**

Sample Text — After: **Corrected**

Translation Fix

Before:
暴露汇总分析数据集）→ Missing opening parentheses

After:
（暴露汇总分析数据集）→ Pair complete

MSD

Benchmark: 20 Clinical Trial Images

Real clinical figures: forest plots, demographic tables, KM curves, clinical posters. EN→ZH.

Vision-Driven: **57 review corrections** · **95 merge groups** · **7/20 quality wins** — at a cost of **+74s** total wall time (+27%)

Dimension	Text-Driven	Vision-Driven	Note
Wall time (20 imgs)	**272s**	346s	+74s (+27%)
Avg per image	**13.6s**	17.3s	Including review step
Review corrections	N/A	**57**	8 of 20 images corrected
Paragraph merge groups	0	**95**	Vision-only capability
CJK width adaptation	N/A	**100%**	All 20 images adjusted
Per-image quality	1 win	**7 wins**	12 ties
API calls / image	1	2	Worker + Review

Pipeline parallelism overlaps OCR with AI calls — wall time only +27% despite higher per-image serial cost.

Pipeline Comparison: PAA Form

Original

PALATABILITY AND ACCEPTANCE ASSESSMENT (PAA)

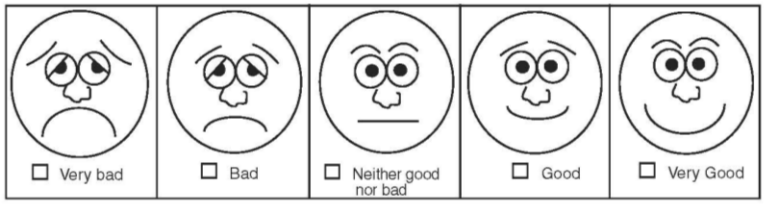
Specify completion date: _____ DD-Mon-YYYY

Please mark one box below to identify the person entering the responses on the questionnaire.

- ☐ The patient
- ☐ The parent/primary caregiver
- ☐ The patient and parent/primary caregiver
- ☐ The health care provider (e.g., physician, nurse, medical assistant or nursing assistant caring for the patient)

This form should be completed at the time of oral medication administration.
Please answer all questions.

1. Please mark the box of one picture below to show how you/the patient feels about the taste of the medication:



2. Please indicate if you/the patient had any of the following problems when taking the medication by mouth:

a. Refusing	☐ Yes	☐ No
b. Spitting Out	☐ Yes	☐ No
c. Throwing Up or Spitting Up	☐ Yes	☐ No
d. Gagging	☐ Yes	☐ No
e. Other, specify	☐ Yes	☐ No

3. Are there any other comments related to the taste of the medication? ☐ Yes ☐ No

Specify comments: _____

I confirm this information is accurate.	Subject's/Caregiver's Initials:	Date:

Text-Driven

适口性和接受度评估（PAA）

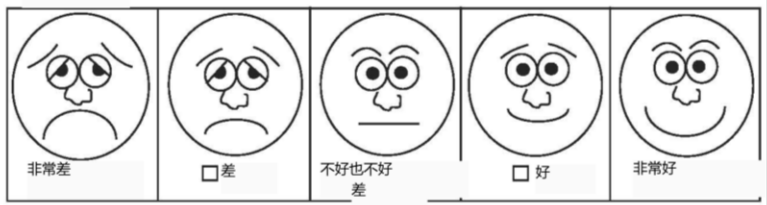
请注明完成日期：_____ 日-月-年

请在下方勾选一个方框，以确定由谁填写本问卷中的回答。

- ☐ 患者
- ☐ 父母/主要照护者
- ☐ 患者和父母/主要照护者
- ☐ The health care provider (e.g., physician, nurse, medical assistant or nursing assistant caring (为患者填写)

本表应在口服给药时填写。
请回答所有问题。

1. 请在下方图片中勾选一个方框，以表示您/患者对该药物味道的感受：



2. 请说明您/患者口服该药物时是否出现以下任何问题：

a. 拒绝服用	是	否
b. 吐出	是	否
c. 呕吐或溢吐	是	否
d. 干呕	是	否
e. 其他，请注明	是	否

3. 是否还有与该药物味道相关的其他意见？

请注明意见：_____

我确认此信息准确无误。	受试者/照护者姓名首字母：	日期：

Vision-Driven

口感及接受度评估（PAA）

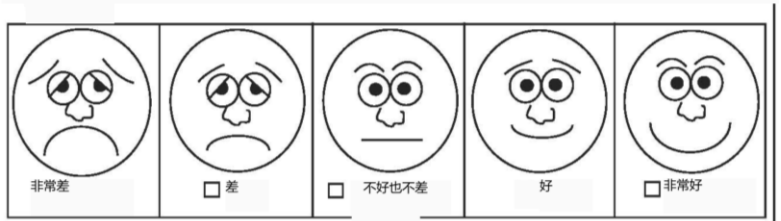
请注明完成日期：_____ 日-月-年

请在下方勾选一个方框，以标识在问卷上填写答案的人员。

- ☐ 患者本人
- ☐ 父母/主要照护者
- ☐ 患者本人及父母/主要照护者
- ☐ 医疗保健提供者（例如，负责照护患者的医生、护士、医疗助理或护理助理）

本表格应在口服给药时填写。请回答所有问题。

1. 请在下方勾选一张图片对应的方框，以表示您/患者对该药物味道的感受：



2. 请说明您/患者在口服药物时是否出现以下任何问题：

a. 拒绝服药	是	否
b. 吐出药物	是	否
c. 呕吐或反流	是	否
d. 作呕	是	否
e. 其他，请说明	是	否

3. 是否有其他关于该药物味道的意见？ 是 否

请说明意见：_____

我确认此信息准确无误。	受试者/照护者姓名缩写：	日期：

OCR miss recovery — where Vision-Driven excels.

MSD

Pipeline Comparison: Clinical Poster

Original

Results

- As the data cut, a total of 55 patients with advanced cervical cancer had been randomized and received at least one dose of Nilvanstomig 10 or 20 mg/kg. Baseline characteristics are shown in Table 1.
- Among the total 22 patients on the 20 mg/kg group who hadn't treated with any prior ICI, the confirmed ORR was 40.9%(Figure 2), and the DCR was 68.2% per IRC's assessments. The mPFS had not yet been reached per the IRC's assessments(Figure 3), but reached 11.01 months per the investigator's assessments.
- Treatment related adverse events (TRAEs) occurred in 46 (83.6%) patients , with 5 (9.1%) being grade 3-4 (one patient each with hypocalcemia, myositis, rash, hypertension and anemia). Serious adverse events (SAEs) occurred in 8 (14.5%) subjects, with only myositis case assessed as related, which also resulted treatment discontinuation. No grade 5 TRAE encountered. Grade 3-4 immune-related adverse events (irAEs) were observed in 4 (7.3%) patients, including myositis, rash, hypertension and anemia (Table 2).
- Nilvanstomig concentration in serum displayed target mediated drug disposition characteristics at lower doses. At doses of 3 mg/kg and above, C_{max} and AUC_{inf} increased in a near dose-proportional manner (Figure 4). Near saturated receptor occupancy on peripheral T cells maintained at least 3 weeks at 3 mg/kg or higher doses, supporting a dosing interval of 3 weeks or longer (Figure 5). Preliminary immunogenicity results indicated low incidence of anti-ZG005 antibodies.

Text-Driven

结果

截至数据截止日，共有55例晚期宫颈癌患者已被随机分组并接受了至少一剂Nilvanstomig 10或20 mg/kg治疗。基线特征见表1。

- 在20 mg/kg组中未接受过任何既往ICI治疗的22例患者中，经IRC评估确认的ORR为40.9%（图2），DCR为68.2%。经IRC评估的mPFS尚未达到（图3），但经研究者评估的mPFS为11.01个月。
- 治疗相关不良事件（TRAEs）发生于46例（83.6%）患者中，其中5例（9.1%）为3-4级（每例分别为低钙血症、肌炎、皮疹、高血压和贫血）。严重不良事件（SAEs）发生于8例（14.5%）受试者中，仅肌炎病例被评估为相关，且导致停止治疗。未发生5级TRAE。3-4级免疫相关不良事件（irAEs）见于4例（7.3%）患者，包括肌炎、皮疹、高血压和贫血（表2）。
- Nilvanstomig血清浓度在较低剂量时表现出靶点介导的药物处置特征。在3 mg/kg及以上剂量时，Cmax和AUCint呈近似剂量比例的方式增加（图4）。外周T细胞上的受体占有率在3 mg/kg或更高剂量下至少维持3周接近饱和，支持3周或更长的给药间隔（图5）。初步免疫原性结果显示抗ZG005抗体发生率低。

Vision-Driven

结果

截至数据截止日，共有55例晚期宫颈癌患者被随机分配并接受了至少一剂Nilvanstomig 10或20 mg/kg治疗。基线特征见表1。

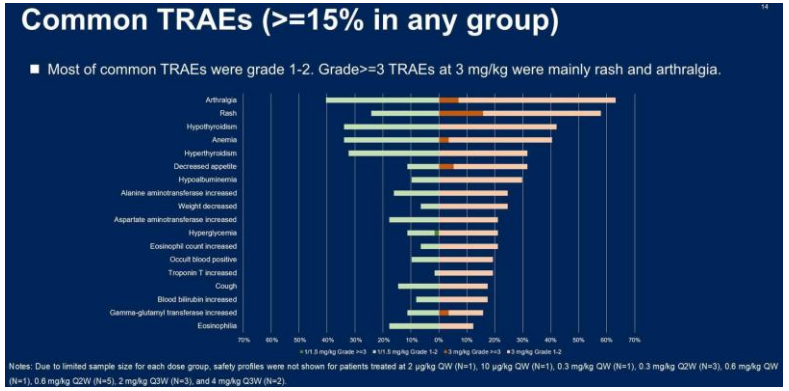
- 在20 mg/kg组中，22例未接受过任何既往ICI治疗的患者中，经IRC评估确认的ORR为40.9%（图2），DCR为68.2%。经IRC评估，mPFS尚未达到（图3），但经研究者评估达到11.01个月。
- 治疗相关不良事件（TRAEs）发生于46例（83.6%）患者中，其中5例（9.1%）为3-4级（各1例分别为低钙血症、肌炎、皮疹、高血压和贫血）。严重不良事件（SAEs）发生于8例（14.5%）受试者中，仅肌炎病例被评估为相关，且导致治疗中断。未发生5级TRAE。3-4级免疫相关不良事件（irAEs）见于4例（7.3%）患者，包括肌炎、皮疹、高血压和贫血（表2）。
- Nilvanstomig血清浓度在较低剂量时表现出靶点介导的药物处置特征。在3 mg/kg及以上剂量时，Cmax和AUCint以近似剂量成比例的方式增加（图4）。在3 mg/kg或更高剂量下，外周T细胞上近饱和的受体占有率维持至少3周，支持3周或更长的给药间隔（图5）。初步免疫原性结果显示抗ZG005抗体发生率低。

MSD

Paragraph merge — Vision-Driven's strongest advantage.

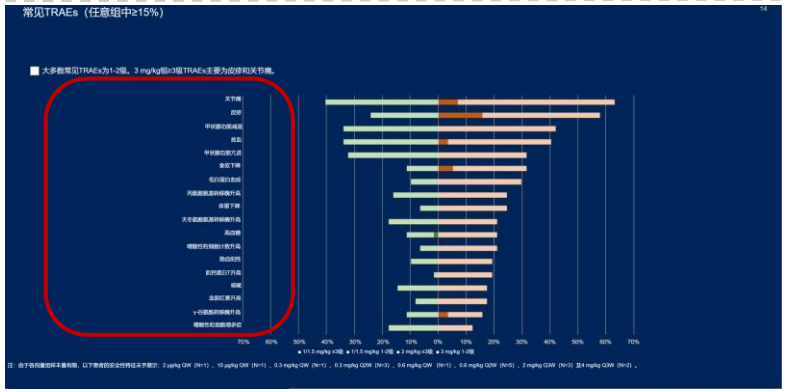
Pipeline Comparison : Dark Background Chart

Original



Text-Driven

Vision-Driven



MSD

Dark background — Vision Worker determines correct text color + right-alignment.

Vision-Driven Pipeline Summary

Strengths

1. Visual context → fewer errors
2. Paragraph merge: denser rendering
3. OCR miss recovery: safety net for missed text
4. Color-aware: AI determines colors from image context
5. Same safety layer: Python Hard Clamp + Canvas Editor unchanged

Trade-offs

1. Full clinical image sent to cloud API (data governance)
2. 2 vision calls/image — slower, pipeline parallelism helps

Both pipelines share the same Python safety bounds and Canvas Editor. Vision-Driven adds visual context and automated QA at the cost of speed and data privacy.

MSD

When to Use Which: A Practitioner's Guide

Text-Driven — Fast & Private

Batch processing

Speed advantage compounds at volume

Data-sensitive figures

Images NEVER leave local server

Standard clinical outputs

Forest plots, KM curves, demographic tables

Vision-Driven — Smart & Thorough

Complex table layouts

AI sees column boundaries and structure

Dense paragraph text

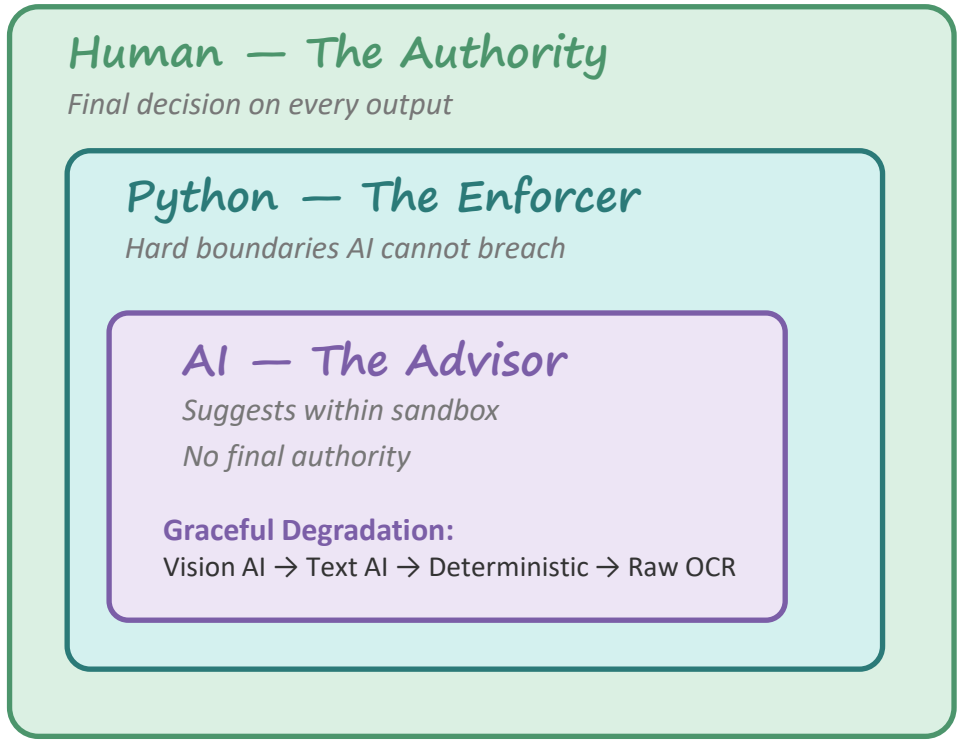
Paragraph merge detection → natural rendering

OCR-challenging content

Missed text auto-recovered by vision model

Production Enough: Text-Driven | Enhanced Mode toggle: Vision-Driven

AI Where It Counts — Not AI Everywhere



Design Rationale — Fit-for-Purpose AI

Same principle as Risk-Based Monitoring (ICH E6 R2):

Token cost is real	→	Code handles everything it can; AI only for judgment calls
Data governance first	→	OCR 100% local; only text/image sent when explicitly chosen
Complexity varies	→	Simple → 1 AI call, Complex → 2 AI calls; not one-size-fits-all
Human time is precious	→	Over-translate by design — deletion is cheap, re-creation is expensive

Trusted Human–AI Collaboration = Humans decide ***where*** and ***how much*** AI to deploy.

MSD

Thank You · 谢谢

Questions Welcome · 欢迎提问

Jianing Wang | Sr. Scientist II, Statistical Programming, BARDS | MSD, China

PHUSE SDE 2026 · Beijing

MSD