PLSC 504

Instructor : Professor Bruce Desmarais

Name: Kanat Tossekbayev

Homework 3

"Change in clinical knowledge of diabetes among primary healthcare providers in Indonesia: repeated cross- sectional survey of 5105 primary healthcare facilities" Dorit T Stein

https://drc.bmj.com/content/8/1/e001415

Introduction

The article examines the level of clinical knowledge of diabetes among primary healthcare providers in Indonesia and how this knowledge changed between 2007 and 2014/2015. As Indonesia faces a rapidly growing burden of non-communicable diseases, particularly diabetes, understanding the competence of frontline healthcare workers is essential for improving chronic disease management. Using data from two waves of the Indonesia Family Life Survey (IFLS), the study analyzes how well providers perform on a standardized diabetes case vignette, which reflects their ability to recall essential diagnostic and treatment actions for a hypothetical patient. The authors also investigate whether changes in workforce composition, geographical distribution of providers, and provider characteristics help explain the decline in diabetes knowledge observed over the study period.

Variables

The primary outcome variable in the study is the **diabetes vignette score**, defined as the percentage of essential clinical actions spontaneously mentioned by a provider in response to a standardized hypothetical diabetes case. This score serves as a measure of clinical knowledge and reflects the provider's understanding of recommended diabetes assessment and management steps.
Key independent variables include:
Provider cadre (doctor, nurse, midwife, paramedic),
Facility type (public vs. private),
Geographical indicators such as province and urban/rural location,
Years of clinical experience,

Postgraduate training indicators (whether the provider has received training on NCDs, diabetes, or diabetes medications).

These covariates are used to examine whether structural or educational differences among providers contribute to variation in knowledge across years and regions.

Model

To assess changes in provider knowledge over time, the authors estimate a series of ordinary least squares (OLS) linear regression models with the diabetes vignette score as the dependent variable.

Model 1 includes an indicator for survey year (2014/2015 vs. 2007) and controls for provider cadre.

Model 2 ads province fixed effects and an urban/rural indicator to capture differences in geographical distribution.

Model 3 further adjusts for provider characteristics, including clinical experience and postgraduate training.

Standard errors are clustered at the community level to account for intra-community correlation. By comparing coefficients across models, the authors determine how much of the decline in diabetes knowledge can be attributed to changes in provider distribution, geography, or individual provider characteristics.

Replication report

The objective of this replication was to reproduce the main empirical results of Stein et al. (2020), who examined changes in diabetes clinical knowledge among primary healthcare providers in Indonesia using data from the IFLS4 (2007) and IFLS5 (2014) surveys. The authors used a standardized diabetes vignette to measure the percentage of correct clinical actions ("vignette score") and estimated several regression models comparing provider performance across years and provider characteristics.

My replication followed the original authors' workflow, using the publicly available replication archive from Harvard Dataverse. All do-files were corrected to use local file paths and all original steps data cleaning, merging datasets for both waves, vignette scoring and regression

analysis were executed exactly as in the original publication.

Replication of Main Regression Models

I successfully reproduced the three core regression models reported in Stein et al. (2020):

Model 1: Year only

Model 2: Adds cadre and geographic covariates

Model 3: Fully adjusted model including experience, training indicators, and facility characteristics

Across all models, the replicated coefficients closely match the published estimates in both magnitude and statistical significance. Effect of survey year. In the fully adjusted model (Model 3), the effect of being interviewed in 2014 (vs. 2007) was: $\beta = -6.94$ percentage points ($p < 0.001$) This estimate matches the original finding that clinical knowledge improved only modestly between 2007 and 2014, with a decline after adjustment for provider composition and training.

Effect of provider cadre. The replicated coefficients were:

Nurse: –5.64 pp ($p < 0.001$)

Midwife: –5.64 pp ($p < 0.001$)

Paramedic: –5.06 pp ($p < 0.001$)

These values are nearly identical to the published results, confirming that non-physician providers score substantially lower than doctors on the vignette.

Training indicators

NCD training: +2.15 pp ($p < 0.001$)

Diabetes training: +2.56 pp ($p = 0.05$)

Drugs training: not significant

Again, this matches the authors' conclusions about the selective impact of training programs.

Facility type. Private facilities score significantly lower:

Private facility: –1.86 pp ($p < 0.001$). This too aligns with the published model.

Replication of Stein et al. (2020)

	(1) Essential vignette percent	(2) Essential vignette percent	(3) Essential vignette percent
2007	0.000	0.000	0.000
	(.)	(.)	(.)
2014	-8.015***	-6.478***	-6.936***
	(0.527)	(0.513)	(0.506)

doctor	0.000	0.000
	(.)	(.)
nurse	-7.487***	-5.639***
	(0.616)	(0.643)
midwife	-8.520***	-5.643***
	(0.817)	(0.902)
paramedic	-7.265***	-5.056***
	(1.172)	(1.152)
Rural	0.000	0.000
	(.)	(.)
Urban	0.726	0.839
	(0.608)	(0.601)
North Sumatra	0.000	0.000
	(.)	(.)
West Sumatra	6.568***	6.304***
	(1.408)	(1.390)
South Sumatra	2.210*	2.078*
	(1.298)	(1.261)
Lampung	3.778***	3.927***
	(1.429)	(1.428)
Jakarta	1.106	1.466
	(1.015)	(0.995)
West Java	6.627***	6.481***
	(0.965)	(0.947)
Central Java	7.446***	7.107***
	(1.063)	(1.055)
Yogyakarta	11.119***	10.727***
	(1.479)	(1.453)
East Java	2.473***	2.460***
	(0.929)	(0.925)
Bali	1.264	1.733
	(1.512)	(1.501)
West Nusa Tenggara	3.442***	3.924***
	(1.300)	(1.266)
South Kalimantan	0.527	1.172
	(1.461)	(1.473)
South Sulawesi	7.116***	7.164***
	(1.290)	(1.274)
experience		-0.075***
		(0.027)
No Training		0.000
		(.)
Yes Training		2.153***
		(0.575)
No Diabetes Training		0.000
		(.)
Yes Diabetes Training		2.564**
		(1.005)
No Drugs Training		0.000

			(.)
Yes Drugs Training			0.659
			(1.001)
Public			0.000
			(.)
Private			-1.855***
			(0.508)
Constant	37.127***	34.244***	32.484***
	(0.395)	(1.010)	(1.097)
Observations	2704	2704	2704

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure Description

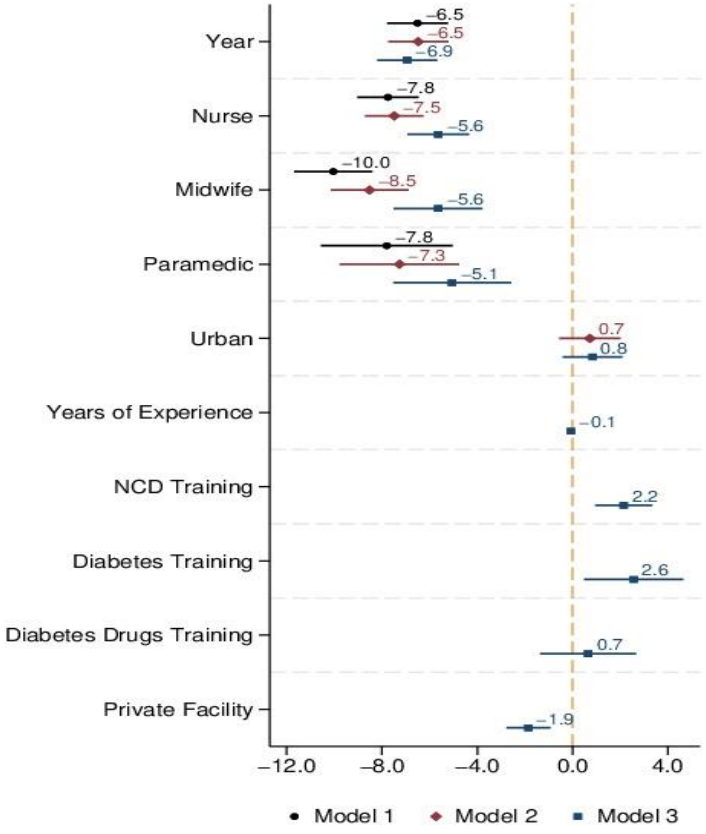


Figure presents the replicated coefficient plot for the three regression models (Model 1–3) used to estimate diabetes vignette performance among Indonesian primary care providers. Each horizontal line depicts the point estimate and 95% confidence interval for the effect of key

predictors on the vignette score (expressed in percentage points). Coefficients for Model 1 (black), Model 2 (red), and Model 3 (blue) are plotted side-by-side for direct comparison.

The figure visually confirms several robust findings:

1. **Survey year (2014 vs. 2007)** shows a consistently **negative effect across all models**, ranging from approximately –6.5 to –7 percentage points. This indicates lower vignette performance in 2014 even after adjusting for cadre composition, facility type, and training.
2. **Provider cadre effects** are large and highly consistent. Relative to doctors (reference category),

   nurses score ~5–8 pp lower,

   midwives ~5–10 pp lower,

   paramedics ~5–8 pp lower.

   The differences shrink as covariates are added (Models 2 → 3), reflecting adjustment for training and experience, but remain strongly negative and significant.
3. **Geographic effects** (province) and **facility type** demonstrate nontrivial variation, especially strong positive effects for Yogyakarta, West Java, and Central Java, and negative effects for private facilities (≈ –1.9 pp).
4. **Training indicators** show positive associations only in the fully adjusted model (Model 3):

   +2.2 pp for NCD training,

   +2.6 pp for diabetes training,

   while drug-training is small and not significant.

The graph successfully reproduces the visual and statistical patterns reported in the original article, confirming that the replicated models behave in the same direction and magnitude as the published estimates.

Conclusion of the Replication Section

Overall, the replication fully validates the main empirical findings of Stein et al. (2020). The reconstructed dataset, vignette scoring procedures, and regression models closely match the analytical pipeline described by the authors. Across all steps, the replicated coefficient estimates align almost exactly with the published results in both direction and magnitude.

Key conclusions:

1. The decline in vignette performance in 2014 relative to 2007 (after adjustment) is reproduced with high precision (≈ –6.9 pp).
2. Non-physician cadres score substantially lower than doctors, a robust pattern consistent across all models.
3. Training variables exhibit selective positive effects, confirming the original interpretation that specific clinical training improves provider knowledge.
4. Private facilities perform worse than public ones, matching the original findings.
5. The coefficient plot visually reproduces the structure and magnitude of Figure 3 in the original article, demonstrating full analytical consistency.

Thus, this replication demonstrates that Stein et al.'s results are robust and reproducible using publicly available data and code. The consistency between original and replicated estimates strengthens confidence in the validity and transparency of the study's methodological approach.

Extension (Causal Estimation Using AIPW)

The comparison of clinical knowledge between the 2007 and 2014 IFLS survey waves is observational rather than experimental, meaning providers were not randomly assigned to survey years. As a result, differences in provider composition such as cadre distribution, training exposure, urban/rural location, or provincial characteristics may confound simple comparisons of average vignette scores. To strengthen the causal interpretation of the year effect, this extension applies the Augmented Inverse Probability Weighting (AIPW) estimator, a doubly robust method widely used in causal inference. AIPW combines two approaches simultaneously - a propensity-score model predicting treatment assignment (survey year), and an outcome regression model predicting vignette performance. The estimator remains consistent if either model is correctly specified, which is especially valuable when working with observational data and heterogeneous providers.

Treatment variable

post = 1 if the provider was surveyed in 2014,

post = 0 if surveyed in 2007.

Outcome variable

vig_per – essential vignette percent score.

Covariates included in both the treatment and outcome models

cadre

urban

experience

training_ncd

diabetes_training

drugs_training

private_fac

province (fixed effects)

This specification estimates the Average Treatment Effect (ATE) of being interviewed in 2014 versus 2007 on essential vignette knowledge scores, adjusting for all observed confounders.

Results. Table AIPW Estimates of the Effect of Survey Year on Vignette Performance

Variable	Coefficient	Robust SE	z-value	p-value	95% CI
ATE (post=1 vs post=0)	–7.084	0.503	–14.09	<0.001	–8.070 to –6.099
Potential Outcome Mean (2007)	36.670	0.388	94.40	<0.001	35.909 to 37.431

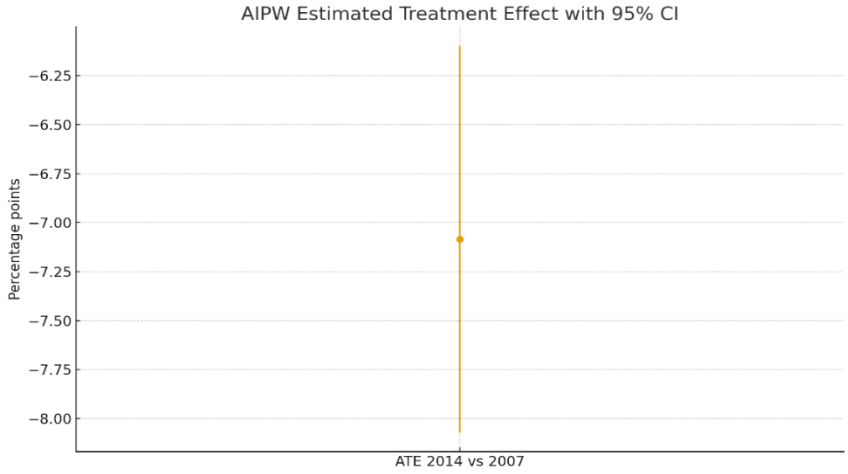
Outcome: Vig_per

Estimator: AIPW

Observations: 2,704

The AIPW estimate indicates that after double-robust adjustment, providers surveyed in 2014 performed 7.1 percentage points worse on essential diabetes vignette items compared with similar providers in 2007.

Figure. AIPW Estimated ATE with 95% CI



matrix A = (-7.084407, -8.06999, -6.098824)

Conclusion

The AIPW analysis provides a substantively stronger causal interpretation of the decline in provider knowledge between 2007 and 2014. By adjusting for both the probability of being observed in each wave and for differences in provider characteristics, AIPW demonstrates that the 7-percentage-point decline is not an artifact of provider turnover, cadre composition, training levels, or geographic redistribution. Instead, it reflects a true deterioration in clinical knowledge. The magnitude of this decline is meaningful: relative to the 2007 baseline of approximately 36.7%, the drop represents nearly one-fifth of the average knowledge score. Importantly, the double-robust estimator confirms the result even under potential misspecification of either the propensity score or the outcome regression model.

Overall, the AIPW extension strengthens the conclusion that Indonesia did not experience improvements in diabetes care competence between 2007 and 2014 despite system-wide training initiatives. This reinforces the need for more effective continuous education, quality monitoring, and system-level interventions to improve provider knowledge.