Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Cognitive screening and driver self-appraisal as predictors of naturalistic hazardous events in cognitively intact older drivers: a prospective cohort study in Japan

Naoki Kamiya
Suzuki Motor Corporation
Satoshi Tanaka
tanakas@hama-med.ac.jp
Hamamatsu University School of Medicine https://orcid.org/0000-0002-6525-682X
Yoshihiro Matsumoto
Suzuki Motor Corporation
Seiji Yamamoto
Hamamatsu University School of Medicine

Research Article

Keywords: older drivers, naturalistic driving, MMSE, self-appraisal, hazardous events, cognitive screening

Posted Date: September 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10826467/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare potential competing interests as follows: Naoki Kamiya and Yoshihiro Matsumoto are employees of Suzuki Motor Corporation. Satoshi Tanaka and Seiji Yamamoto were financially supported by grants from Suzuki Motor Corporation.

Abstract

Background

Although cognitive screening and self-appraisal have been examined in relation to driving performance on road tests and simulators, whether they correspond to naturalistically recorded driving behavior in cognitively intact older adults remains unclear. We examined associations of the Mini-Mental State Examination (MMSE) score and three self-appraisal indicators with hazardous driving events in a naturalistic cohort.

Methods

In a prospective cohort, 192 cognitively intact drivers aged 60 to 75 years (mean MMSE 29.3, SD 1.0) in Hamamatsu, Japan, had telematics installed in their vehicles for three months. Rapid acceleration, hard braking, and sharp steering were recorded automatically at a fixed kinematic threshold. Negative binomial regression, with log driving time as an offset and adjustment for age, sex, and education, tested associations with MMSE score and with three self-appraisal indicators (self-rated driving ability and self-perceived cognitive decline while driving and in general). Equivalence analysis quantified the precision of null results.

Results

MMSE score was not associated with the rate of any event type (incidence rate ratios [IRRs] 0.845 to 1.089; all corrected $p \geq .224$). Equivalence analysis constrained the effect to IRRs below about 1.2 to 1.4, excluding moderate-to-large associations. None of the three self-appraisal indicators was associated with any event rate (all corrected $p = 1.00$). Adjusting for advanced driver-assistance systems did not alter the estimates.

Conclusions

In cognitively intact older drivers, neither MMSE score nor brief self-appraisal predicted naturalistic hazardous event rates. Global cognitive screening does not appear to capture real-world driving risk in this population.

1. Introduction

Population aging has placed older driver safety at the center of clinical, policy, and public-health discussion across high-income societies (Li et al., 2003). In Japan, adults aged 65 years and older accounted for 29.3% of the population in 2024, the highest proportion of any country, and the number of older license holders continues to rise (Cabinet Office, 2025). In response, Japan introduced mandatory

cognitive testing for older drivers at license renewal, but subsequent evaluations indicate that the policy has had a limited and mixed effect on the overall safety of older road users (Ichikawa et al., 2020; Inada et al., 2023). Therefore, identifying which assessments reliably index actual driving risk in this population carries both clinical and policy weight, and the recent expansion of in-vehicle sensing technology has begun to enable observation of older drivers' real-world behavior at scales not previously available (Dingus et al., 2006, 2016; Eby et al., 2019).

Among brief cognitive instruments, the Mini-Mental State Examination (MMSE) (Folstein et al., 1975) is the best known and most frequently used screening measure of global cognitive status across clinical, research, and community settings (Arevalo-Rodriguez et al., 2015; Tsoi et al., 2015), and it has been applied extensively in older-driver research (Mathias and Lucas, 2009; Crizzle et al., 2012). It is therefore commonly used as the cognitive reference point in studies of older-driver safety.

Within the cognitively intact range, however, the MMSE's predictive validity for driving is not fully established, since the positive evidence comes largely from samples that also include cognitively impaired drivers. For example, higher MMSE scores have been associated with safer driving in samples that combine cognitively intact and impaired older drivers, both on standardized road tests and in naturalistic recordings (Reger et al., 2004; Davis et al., 2012). However, a meta-analytic review restricted to healthy older adults found no correlation between the MMSE and driving performance, whereas domain-specific measures such as the Trail Making Test Part B showed small-to-moderate associations (Stefanidis et al., 2023). Systematic-review evidence similarly points to specific domains, particularly executive function and complex attention, rather than global status, as most relevant to driving (Depestele et al., 2020). This evidence, however, derives almost entirely from standardized on-road tests and simulators, controlled settings that may not reflect the self-paced demands of everyday driving.

A second question concerns how older drivers perceive their own functioning, and whether such self-appraisal corresponds to objectively recorded risk (Anstey et al., 2005; Horrey et al., 2015). Self-appraisal can take more than one form: drivers may judge their overall driving ability (Marottoli and Richardson, 1998), or report awareness of cognitive decline, either specifically while driving or in everyday life (Jessen et al., 2014). These indicators have been studied in relation to driving cessation and self-regulation (Baldock et al., 2006; Molnar and Eby, 2008; Ross et al., 2009). Although a recent large-scale Japanese cross-sectional study has linked subjective memory concerns to self-reported collisions in older drivers (Kurita et al., 2024), whether such subjective indicators correspond to objectively measured hazardous events in naturalistic settings has not been established.

The present study addresses these gaps in a prospective cohort of cognitively intact older drivers (n = 192). We examined whether (1) MMSE score and (2) three self-appraisal indicators (self-rated driving ability, self-perceived cognitive decline while driving, and self-perceived cognitive decline in general) are associated with the rate of hazardous driving events (rapid acceleration, hard braking, and sharp steering) recorded over three months. The MMSE analysis extends to a sensor-based naturalistic setting

a question previously examined only with on-road and simulator measures, and assessing objective and subjective predictors in one cohort allows their convergence to be evaluated.

2. Methods

2.1 Study design and setting

This was a prospective observational cohort study using a naturalistic driving methodology. Participants were recruited between September 2021 and March 2023 in Hamamatsu City, Shizuoka, Japan. The study was approved by the Institutional Review Board of Hamamatsu University School of Medicine (approval number 21–130).

2.2 Participants

Participants were recruited through three channels: (1) community-based venues for older adults, (2) local senior clubs, and (3) mandatory driver training courses for older adults held at a local driving school, where printed study materials were distributed and on-site oral recruitment was conducted. Eligible participants were community-dwelling older adults aged 60 to 75 years who drove on average at least three days per week and held a valid driver's license. Individuals were excluded if their MMSE score was 23 or below. Sex was recorded as self-reported biological sex (male/female) at the baseline interview; no measure of gender identity was collected. Written informed consent was obtained from all participants before enrolment. The target sample size was set at 200, representing the maximum number feasible within the operational capacity of the study rather than being based on a formal a priori power calculation. This sample is larger than that of the largest previous naturalistic driving study to relate the MMSE to a recorded driving outcome in older drivers (Davis et al., 2012, n = 103).

2.3 Naturalistic driving data collection

Driving data were collected using an onboard telematics device (SD4-LTE, SmartDrive Co., Ltd., Tokyo, Japan), which was installed in the cigarette-lighter socket of each participant's personal vehicle and remained in place throughout the three-month observation period. The device continuously recorded tri-axial acceleration and GPS position during each trip, and data were transmitted wirelessly to a secure server on a trip-by-trip basis. A trip was operationally defined as the interval from engine ignition to engine shut-off. Participants were instructed to install the device on a vehicle that they themselves primarily drove; when a vehicle was shared with family members, they were asked to remove the device before others used it.

Hazardous driving events were operationally defined as rapid acceleration (longitudinal positive), hard braking (longitudinal negative), and sharp steering (lateral), each detected when the corresponding acceleration reached 0.4 G or greater, equivalent to a one-second velocity change of approximately 14 km/h. This threshold aligns with the operational definition used in large-scale naturalistic driving studies (Longitudinal Research on Aging Drivers: LongROAD) of older adult drivers (e.g., Li et al., 2017; Liu et al.,

2024). The number of each event type accumulated over the observation period served as the outcome variable.

2.4 Baseline covariates and other recorded variables

At the baseline interview, additional information was recorded for use as covariates or for descriptive characterization of the sample: demographic variables (age, sex, years of education), driving experience (weekly driving frequency, years since license acquisition, annual driving distance), and the presence of advanced driver-assistance systems (ADAS).

2.5 Cognitive assessment

The Mini-Mental State Examination-Japanese (MMSE-J; Sugishita et al., 2018) was administered at the baseline session, before device installation. The MMSE is a 30-point screening instrument assessing orientation, registration, attention, recall, language, and visuospatial ability.

2.6 Subjective self-appraisal indicators

Three items assessing subjective self-appraisal were administered orally by a researcher as part of a baseline questionnaire. The first item, self-rated driving ability, asked participants to rate their current driving ability on a three-point ordinal scale (poor / neither / good). The second item, self-perceived cognitive decline while driving, asked whether participants had noticed any decline in their cognitive function specifically while driving (yes / no). The third, self-perceived cognitive decline in general, asked the same question about everyday life more broadly (yes / no). The two cognitive-decline items were retained as distinct indicators because awareness of decline while driving and awareness of decline in general need not coincide, and each may bear a different relationship to driving behavior.

2.7 Data analysis

2.7.1 Trip-level data filtering

Three trip-level filters were applied before analysis. First, trips shorter than approximately 2 km were excluded, as such trips provide too little continuous driving to characterize driving behavior reliably. Second, trips in which idling time exceeded 50% of total trip duration were excluded, to remove trips dominated by stationary periods. Third, trips with an average speed greater than 35 km/h or a total distance greater than 60 km were excluded, to remove long-distance expressway trips in which hazardous events are unlikely to occur. The number of trips per participant was not used as the unit of analysis; instead, total driving time accumulated across all retained trips served as the exposure metric, entered as an offset in the negative binomial models (see Section 2.7.3).

2.7.2 Descriptive statistics

Descriptive statistics were computed for all variables. Continuous variables are summarized as mean ± SD; when distributions were strongly right-skewed (driving exposure and event counts), the median with

IQR is reported instead of, or in addition to, the mean. Categorical variables are reported as counts and percentages. Unless otherwise stated, all inferential tests were two-sided, with $\alpha = .05$.

2.7.3 Modelling of hazardous event rates

Hazardous driving events were analyzed as counts. Because each participant contributed a different total driving exposure, all counts were modeled with the natural logarithm of driving time (hours) entered as an offset, so that the effective outcome was the event rate per unit driving time. The three event types (rapid acceleration, hard braking, and sharp steering) were analyzed separately.

Before fitting the regression models, bivariate associations among the continuous variables (the three event rates, mean speed, and MMSE score) were examined with Spearman rank correlations as preliminary descriptive analyses.

Overdispersion of each event count was assessed with the variance-to-mean ratio (VMR), calculated as the empirical variance divided by the mean. Because all three event counts were overdispersed, we fitted negative binomial regression models (NB2 parameterization, log link), in which the dispersion parameter was estimated from the data rather than fixed. This modeling approach has precedent in prior naturalistic driving studies of older drivers (Keay et al., 2013; Liu et al., 2024; Eby et al., 2019). For each outcome, the negative binomial model was compared with a Poisson model with the same linear predictor, using the Akaike Information Criterion (AIC) and a likelihood-ratio test.

All models adjusted for the same three a priori covariates: age (years), sex (male as the reference category), and years of education. These were selected as theoretically motivated confounders that are plausibly associated with the outcome and not presumed to lie on the causal pathway between the predictor of interest and hazardous events.

For the first research question, MMSE score was entered as a continuous predictor in three negative binomial models, one per event type. For the second, each of the three self-appraisal indicators was tested against the three event types, yielding nine models: self-rated driving ability (three-level ordinal, poor / neither / good), self-perceived cognitive decline while driving (yes / no), and self-perceived cognitive decline in general (yes / no). Self-rated driving ability, a three-level categorical variable, was evaluated with a single Wald test of the joint null hypothesis that both non-reference coefficients equalled zero (2 degrees of freedom), so that it counted as one test rather than two.

Associations are reported as incidence rate ratios (IRRs) with 95% confidence intervals. To control the family-wise error rate, a Bonferroni correction was applied separately within each research question. Both uncorrected and corrected p-values are reported. The first research question concerned associations whose absence is itself of interest. For the MMSE null results, we therefore quantified their precision using two descriptors that do not depend on the significance threshold. The first was the minimum effect detectable with 80% power, given the observed exposure and standard errors. The second was the equivalence margin. We defined the equivalence margin as the narrowest incidence-rate ratio within which the two-sided 90% confidence interval was wholly contained (Lakens, 2017; Lakens et

al., 2018). We did not designate any effect as negligible in advance. Instead, we report the achieved bound, so that readers can judge it against whatever threshold they consider meaningful. Analyses were conducted in R version 4.5.1 using the MASS package (glm.nb) for negative binomial regression; the joint Wald test for self-rated driving ability was computed directly from the fitted model's coefficient vector and covariance matrix. Equivalence margins, minimum detectable effects, and the robustness analyses (Section 3.3) were computed in the same R environment.

2.7.4 Sensitivity analyses

As a sensitivity analysis, each model was refitted as an ordinary least-squares regression on the log-transformed event rate [log((count + 0.5) / driving hours)]. Any predictor reaching nominal significance was further examined with leave-one-out influence diagnostics.

As an additional sensitivity analysis, all twelve models were refitted with the presence of advanced driver-assistance systems (ADAS), extracted from baseline vehicle records and dichotomized as present versus absent, added as a covariate.

3. Results

3.1 Participant characteristics

Of the 200 enrolled participants, five were excluded as ineligible and three withdrew during the observation period for personal reasons, leaving 192 who completed the full three-month data collection and were included in all analyses. No motor vehicle crashes or other adverse events were reported. Participant characteristics are summarized in Table 1. The sample was predominantly male (121 participants, 63.0%), with a mean age of 69.1 years (SD = 4.4). Participants were experienced drivers (mean driving history 47.1 years, SD 6.5), drove a mean of 5.9 days per week (SD 1.5; 58.3% drove daily), and reported a median annual driving distance of 6,000 km (IQR 4,000–10,000). MMSE scores ranged from 24 to 30 (mean 29.3, SD 1.0; median 30). Of the 192 participants, 115 (59.9%) scored the maximum and 94.8% scored ≥ 28. For self-rated driving ability, 40 participants (20.8%) rated themselves as poor, 95 (49.5%) as neither, and 57 (29.7%) as good. Eighty-nine participants (46.4%) reported having noticed a decline in their cognitive function while driving, and 138 (71.9%) reported the same in everyday life more generally.

After application of the trip-level filters, 39,323 of 61,677 recorded trips (63.8%) were retained for analysis (median 191 trips per participant, IQR 129–265). Participants accumulated a median total driving distance of 1,079 km (IQR 692–1,522) and a median total driving time of 52.2 hours (IQR 35.0–77.1) over the three-month observation period.

Table 1
Participant characteristics (n = 192).

Characteristic	Value
Age, years, mean (SD)	69.1 (4.4)
Sex, male, n (%)	121 (63.0%)
Education, years, mean (SD)	13.6 (2.2)
MMSE score, mean (SD)	29.3 (1.0)
Driving frequency, days/week, mean (SD)	5.9 (1.5)
Driving history, years, mean (SD)	47.1 (6.5)
Annual driving distance (self-reported), km, median (IQR)	6,000 (4,000–10,000)
Self-rated driving ability	
Poor, n (%)	40 (20.8%)
Neither, n (%)	95 (49.5%)
Good, n (%)	57 (29.7%)
Self-perceived cognitive decline while driving (yes), n (%)	89 (46.4%)
Self-perceived cognitive decline in general (yes), n (%)	138 (71.9%)
Observation-period driving distance, km, median (IQR)	1,079 (692–1,522)
Observation-period driving time, h, median (IQR)	52.2 (35.0–77.1)

IQR = interquartile range; MMSE = Mini-Mental State Examination; SD = standard deviation.

3.2 Hazardous driving events

Across the 192 participants, the total number of events was 1,163 for rapid acceleration, 11,785 for hard braking, and 3,964 for sharp steering. Hard braking was the most frequent event type (median 39.0, IQR 17.0–75.2), followed by sharp steering (median 10.0, IQR 3.0–26.0) and rapid acceleration (median 3.0, IQR 1.0–7.0). All three event counts were strongly right-skewed and severely overdispersed, with variance-to-mean ratios ranging from approximately 12 to 97. Consistent with this, negative binomial models were decisively favored over Poisson models for every event type (ΔAIC = 931.0, 10,090.4, and 4,822.6 for rapid acceleration, hard braking, and sharp steering, respectively; all likelihood-ratio $p < .001$), confirming the negative binomial specification used throughout. The three event rates were moderately to strongly intercorrelated (Spearman rho = .49–.62). Mean speed was positively correlated with all three event rates (rho = .21–.33, all $p < .01$), whereas MMSE score showed no significant simple correlation with any event rate (all $|rho| \leq .13$, all $p \geq .079$).

3.3 MMSE scores and hazardous driving events

In negative binomial regression models adjusted for age, sex, and education, MMSE score was not significantly associated with the rate of any hazardous event type after correction for multiple comparisons (Table 2; Fig. 1). The incidence rate ratio per one-point increase in MMSE was 0.857 (95% CI 0.724–1.015) for rapid acceleration, 1.089 (0.943–1.256) for hard braking, and 0.845 (0.697–1.024) for sharp steering. None reached the Bonferroni-corrected threshold (α = .0167): the uncorrected p-values were .075, .245, and .087, respectively, all exceeding .05/3 after correction (p_Bonf = .224, .736, and .260). Expressed per one-point increase in MMSE, the equivalence margins were incidence-rate ratios of 1.23 for hard braking, 1.34 for rapid acceleration, and 1.39 for sharp steering. With 80% power the study could detect per-point ratios as small as 1.27 for hard braking, 1.32 for rapid acceleration, and 1.37 for sharp steering, after correction for multiple comparisons. A one-point change in MMSE corresponded to 0.97 standard deviations, so per-point and per-standard-deviation effects were essentially identical.

The two outcomes with uncorrected p-values below .10 (rapid acceleration and sharp steering) were examined further and judged not to reflect a substantive association. Refitting these models as ordinary least-squares regressions on a continuity-corrected log event rate again yielded clearly non-significant associations for MMSE (p = .50 for rapid acceleration and p = .33 for sharp steering), and this conclusion did not depend on how zero counts were handled. Leave-one-out diagnostics confirmed that the borderline values were not robust: across single-observation deletions the uncorrected p-value ranged from .03 to .40 for rapid acceleration and from .04 to .26 for sharp steering. The direction of the point estimates was also inconsistent across event types (IRR below 1 for acceleration and steering but above 1 for braking), providing no coherent pattern of a protective or harmful effect of higher MMSE scores.

3.4 Subjective self-appraisal and hazardous driving events

None of the three self-appraisal indicators was significantly associated with any hazardous event rate, either before or after correction for multiple comparisons (Table 2; Fig. 2). For self-rated driving ability, the joint Wald test comparing the three ability categories was non-significant for every event type (p = .399 for rapid acceleration, .439 for hard braking, and .729 for sharp steering). For self-perceived cognitive decline while driving, the incidence rate ratios for participants reporting decline, relative to those reporting none, ranged from 1.033 to 1.089 across the three event types, with all uncorrected p-values ≥ .54. For self-perceived cognitive decline in general, the incidence rate ratios ranged from 0.905 to 1.362, with the smallest uncorrected p-value being .144 (sharp steering). Across all nine tests, every uncorrected p-value was ≥ .144, so all Bonferroni-corrected p-values equaled 1.00. Effect estimates were small, and their confidence intervals all included the null value of 1.

Table 2
Negative binomial regression of hazardous event rates on MMSE and subjective self-appraisal indicators (n = 192).

Predictor	Rapid acceleration	Hard braking	Sharp steering
Outcome rate, events/h, median (IQR)	0.06 (0.02–0.14)	0.78 (0.32–1.34)	0.19 (0.06–0.50)
MMSE (per + 1 point)			
IRR [95% CI]	0.857 [0.724, 1.015]	1.089 [0.943, 1.256]	0.845 [0.697, 1.024]
p_raw / p_Bonf	.075 / .224	.245 / .736	.087 / .260
Self-rated driving ability (joint Wald, 2 df)			
p_raw / p_Bonf	.399 / 1.00	.439 / 1.00	.729 / 1.00
Cognitive decline – while driving (yes vs no)			
IRR [95% CI]	1.056 [0.758, 1.472]	1.089 [0.827, 1.435]	1.033 [0.711, 1.502]
p_raw / p_Bonf	.746 / 1.00	.543 / 1.00	.865 / 1.00
Cognitive decline – general (yes vs no)			
IRR [95% CI]	0.905 [0.628, 1.303]	1.117 [0.822, 1.517]	1.362 [0.900, 2.060]
p_raw / p_Bonf	.591 / 1.00	.479 / 1.00	.144 / 1.00

All models adjusted for age, sex (male reference), and years of education, with log(driving time in hours) as an offset. CI = confidence interval; IRR = incidence rate ratio; IQR = interquartile range; MMSE = Mini-Mental State Examination. p_raw = uncorrected p-value; p_Bonf = Bonferroni-corrected p-value (within research question 1, 3 tests; within research question 2, 9 tests). Self-rated driving ability (poor / neither / good) was evaluated with a single joint Wald test, so no single IRR is reported.

3.5 Sensitivity analysis with ADAS adjustment

Of the 192 participants, 103 (53.6%) drove vehicles equipped with ADAS. Event rates did not differ between groups (Mann–Whitney U; all $p \geq .209$). Adding ADAS as a covariate left every focal estimate essentially unchanged (absolute change in IRR ≤ 0.005; absolute change in $p \leq 0.048$ across all twelve models), and ADAS itself was not associated with any event rate (IRRs 0.948–1.074; all $p \geq .618$). Tests

of MMSE × ADAS interaction were non-significant for all three event types ($p \geq .097$). The conclusions of the main analyses are therefore robust to the presence of ADAS.

4. Discussion

In this prospective naturalistic driving study of cognitively intact older drivers, neither cognitive screening with the MMSE nor subjective self-appraisal predicted the rate of hazardous driving events over three months. To our knowledge, this is the first study to relate the MMSE on its own to naturalistic hazardous events within an exclusively cognitively intact cohort, and the first to relate multiple forms of driver self-appraisal to objectively recorded hazardous events. Earlier naturalistic studies addressed adjacent questions, mapping neuropsychological-domain composites onto driving outcomes (Aschenbrenner et al., 2022), using the MMSE only for enrolment (Seelye et al., 2017), or examining it in mixed samples of cognitively intact and impaired drivers (Davis et al., 2012). The only prior naturalistic study to relate a self-referential index to recorded events indexed cognitive state rather than driving ability (Guo et al., 2015a, 2015b; considered below). With 192 cognitively intact drivers, the present study evaluates in parallel the global screen most relevant to clinical practice and three self-appraisal constructs against an ecologically valid behavioral outcome.

The present MMSE result is best read as a bounded null rather than a simple non-finding. Equivalence analysis constrained the per-point effect on every event rate to an incidence-rate ratio below roughly 1.2 to 1.4 (Section 3.3). In other words, the data exclude the moderate-to-large associations seen once impaired drivers are included, while remaining agnostic about small effects. Consistent with this, naturalistic risky-driving behaviors reliably distinguish mild Alzheimer's disease from healthy controls (Davis et al., 2018). Within the cognitively intact range, then, the MMSE does not appear to differentiate kinematic risk to a degree of practical use. We consider that the ceiling effect is not just a limitation here. It is itself one of the findings. In this cohort 94.8% scored 28 or above, and the central half of the sample spanned a single MMSE point. A screen with so little spread can barely separate higher-risk from lower-risk drivers inside the group it is applied to. Accordingly, the per-point equivalence bounds reported above (≈ 1.2–1.4) exclude only large per-point effects. They cannot speak to gradients the restricted range does not express. The pattern is consistent with meta-analytic evidence that the MMSE does not track driving performance in healthy older adults. Stefanidis et al. (2023) attributed that absence to the instrument's ceiling, noting that domain-specific measures of executive function and complex attention retained small-to-moderate associations (Depestele et al., 2020).

This does not mean the MMSE is unrelated to driving in general. Such measures have been associated with driving once the cognitive range extends beyond the intact (Reger et al., 2004; Mathias and Lucas, 2009). Two cases show where that signal lies. The first is the impaired end of cognition. Among drivers with cognitive or visual dysfunction, a global cognitive composite tracked kinematic events (Merickel et al., 2019). The second is the more demanding setting of a structured road test. Davis et al. (2012) found that the MMSE related only weakly to everyday naturalistic errors, but more strongly to errors on the road

test. Thus, the MMSE may carry predictive signal where cognition is impaired or where the task is demanding, but not within the intact range under self-paced everyday driving.

The null self-appraisal result warrants separate consideration, because the self-appraisal literature has developed largely without objective real-world referents. Older drivers consistently rate their ability as average or better, and self-rating was unrelated to on-road performance and adverse-event history in earlier cohorts (Marottoli and Richardson, 1998; Horswill et al., 2013). Notably, the single previous study to relate a continuously recorded naturalistic outcome to a self-referential index reported the opposite of a null: a larger discrepancy between self-estimated and objectively tested cognitive status was associated with a higher rate of high-g-force events (Guo et al., 2015a). That study indexed cognitive status rather than driving ability, restricted this association to high-g-force rather than crash outcomes, and rested on a small sample. Therefore, the present result, obtained in a larger cohort with explicit self-appraisal measures, does not overturn it but indicates that such an association might not be detectable with single-item self-appraisal indicators in cognitively intact drivers.

Compensatory self-regulation may also help explain the absence of association. Drivers who perceive reduced capacity may avoid demanding situations rather than generate more hazardous events. Molnar and Eby (2008) found self-regulation to be associated with visual but not cognitive impairment, which might indicate that compensatory avoidance is triggered more readily by sensory than by cognitive decline. If so, the mechanism would be expected to fit self-rated driving ability more closely than awareness of cognitive decline, and might apply only weakly to the cognitive indicators that predominate here. Such regulation would in any case be expected to operate at the strategic level of trip planning and avoidance rather than the operational level of vehicle control (Michon, 1985; Hatakka et al., 2002). The event-rate outcome would not be expected to capture adjustments at that level.

For policy, the implication is narrow but concrete. Within the cognitively intact range, a brief global cognitive score did not separate older drivers who generated more hazardous events from those who generated fewer. This is consistent with national evaluations showing only a limited and mixed effect of the renewal cognitive-screening policy on the overall safety of older road users (Ichikawa et al., 2020; Inada et al., 2023). Identifying risk among drivers who pass such screening may require driving-specific assessment or direct naturalistic observation rather than a global score alone.

This study has several limitations. First, the design could not exclude small effects within the cognitively intact range: two of the three MMSE estimates (for rapid acceleration and sharp steering) reached uncorrected p-values just below .10, and although neither survived the alternative analytic specification or correction for multiple comparisons (Section 3.3), a small true association cannot be ruled out. Second, the three-month window, although longer than a single road test, may be too short to capture rare safety-critical events such as crashes, so the findings speak to routine hazardous maneuvers rather than crash risk. Third, the trip-level filters removed 36.2% of the recorded trips, including stationary trips and long-distance high-speed trips, so the exposure analyzed represents routine local driving rather than the full range of driving contexts. Fourth, each self-appraisal construct was measured with a single brief

item rather than a validated multi-item scale, which may have reduced sensitivity to graded self-awareness (Marottoli and Richardson, 1998; George et al., 2007). Fifth, this was an observational study with a single wave of predictor measurement, so the analysis is associational and cannot establish a causal effect. Finally, sex was included only as an adjustment covariate; the study was not designed or powered to examine sex differences in the associations between cognitive or self-appraisal measures and hazardous events, and such differences cannot be excluded.

5. Conclusions

In this naturalistic cohort of cognitively intact older drivers, neither MMSE score nor brief self-appraisal predicted the rate of hazardous driving events. Global cognitive screening does not appear to capture real-world driving risk in this population. Identifying at-risk drivers in this group may instead require driving-specific assessment or continuous telematics.

Declarations

CRediT authorship contribution statement

Naoki Kamiya: Conceptualization, Methodology, Investigation, Data curation, Writing – review & editing. Satoshi Tanaka: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Supervision. Yoshihiro Matsumoto: Conceptualization, Methodology, Investigation, Writing – review & editing, Project administration, Funding acquisition. Seiji Yamamoto: Conceptualization, Methodology, Writing – review & editing.

Declaration of competing interest

Naoki Kamiya and Yoshihiro Matsumoto are employees of Suzuki Motor Corporation. Satoshi Tanaka and Seiji Yamamoto were financially supported by grants from Suzuki Motor Corporation.

Funding

This work was financially supported by Suzuki Motor Corporation under a collaborative research agreement with Hamamatsu University School of Medicine. Through its co-authors employed by the company, the funder contributed to the study design and data acquisition. The statistical analysis, interpretation of the results, and drafting of the manuscript were carried out by the academic authors. As an organization, Suzuki Motor Corporation reviewed the manuscript prior to submission solely for the protection of confidential and proprietary information; this review did not alter the scientific content of the manuscript or the decision to submit it for publication.

Ethics statement

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of Hamamatsu University School of Medicine (approval number 21-130).

Written informed consent was obtained from all participants before enrolment.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the authors used Claude Opus 5 and Fable 5 (Anthropic) in order to (i) generate and debug portions of the R code used for the statistical analysis, and (ii) improve the language and readability of the manuscript text. After using these tools, the authors reviewed, tested, and edited the output as needed and take full responsibility for the content of the published article and for the integrity of the analysis.

Acknowledgments

This study was conducted as a social demonstration project under the Hamamatsu Wellness Project, a public–private initiative led by the Hamamatsu City, Shizuoka, Japan.

References

1. Anstey KJ, Wood J, Lord S, Walker JG (2005) Cognitive, sensory and physical factors enabling driving safety in older adults. Clin Psychol Rev 25:45–65. https://doi.org/10.1016/j.cpr.2004.07.008
2. Arevalo-Rodriguez I, Smailagic N, Roqué i Figuls M, Ciapponi A, Sanchez-Perez E, Giannakou A, Pedraza OL, Cosp B, Cullum X, S (2015) Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). https://doi.org/10.1002/14651858.CD010783.pub2. Cochrane Database of Systematic Reviews, CD010783
3. Aschenbrenner AJ, Murphy SA, Doherty JM, Johnson AM, Bayat S, Walker A, Peña Y, Hassenstab J, Morris JC, Babulal GM (2022) Neuropsychological correlates of changes in driving behavior among clinically healthy older adults. Journals Gerontology: Ser B 77:1769–1778. https://doi.org/10.1093/geronb/gbac101
4. Baldock MRJ, Mathias JL, McLean AJ, Berndt A (2006) Self-regulation of driving and its relationship to driving ability among older adults. Accid Anal Prev 38:1038–1045. https://doi.org/10.1016/j.aap.2006.04.016
5. Cabinet Office, Government of Japan (2025) Annual Report on the Ageing Society [Reiwa 7-nen ban Kōrei Shakai Hakusho]. Cabinet Office, Tokyo. https://www8.cao.go.jp/kourei/whitepaper/w-2025/zenbun/07pdf_index.html (accessed 10 June 2026)
6. Crizzle AM, Classen S, Bédard M, Lanford D, Winter S (2012) MMSE as a predictor of on-road driving performance in community dwelling older drivers. Accid Anal Prev 49:287–292.

https://doi.org/10.1016/j.aap.2012.02.003

7. Davis JD, Papandonatos GD, Miller LA, Hewitt SD, Festa EK, Heindel WC, Ott BR (2012) Road test and naturalistic driving performance in healthy and cognitively impaired older adults: does environment matter? J Am Geriatr Soc 60:2056–2062. https://doi.org/10.1111/j.1532-5415.2012.04206.x
8. Davis JD, Wang S, Festa EK, Luo G, Moharrer M, Bernier J, Ott BR (2018) Detection of risky driving behaviors in the naturalistic environment in healthy older adults and mild Alzheimer's disease. Geriatrics 3:13. https://doi.org/10.3390/geriatrics3020013
9. Depestele S, Ross V, Verstraelen S, Brijs K, Brijs T, van Dun K, Meesen R (2020) The impact of cognitive functioning on driving performance of older persons in comparison to younger age groups: a systematic review. Transp Res Part F: Traffic Psychol Behav 73:433–452. https://doi.org/10.1016/j.trf.2020.07.009
10. Dingus TA, Klauer SG, Neale VL, Petersen A, Lee SE, Sudweeks J, Perez MA, Hankey J, Ramsey D, Gupta S, Bucher C, Doerzaph ZR, Jermeland J, Knipling RR (2006) The 100-Car Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment. Report No. DOT HS 810 593. National Highway Traffic Safety Administration, Washington, DC
11. Dingus TA, Guo F, Lee S, Antin JF, Perez M, Buchanan-King M, Hankey J (2016) Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences 113, 2636–2641. https://doi.org/10.1073/pnas.1513271113
12. Eby DW, Molnar LJ, Zakrajsek JS, Ryan LH, Zanier N, St. Louis RM, Stanciu SC, LeBlanc D, Bogard S, Kostyniuk LP, Smith J, Yung R, Nyquist L, Betz ME, DiGuiseppi C, Jones V, Li G, Mielenz TJ, Strogatz D on behalf of the AAA LongROAD Research Team, 2019. Factors related to rapid deceleration events among a large cohort of older drivers. Transp Res Part F: Traffic Psychol Behav 66, 76–86. https://doi.org/10.1016/j.trf.2019.08.021
13. Folstein MF, Folstein SE, McHugh PR (1975) Mini-mental state: a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198. https://doi.org/10.1016/0022-3956(75)90026-6
14. George S, Clark M, Crotty M (2007) Development of the Adelaide Driving Self-Efficacy Scale. Clin Rehabil 21:56–61. https://doi.org/10.1177/0269215506071284
15. Guo F, Fang Y, Antin JF (2015a) Evaluation of older driver fitness-to-drive metrics and driving risk using naturalistic driving study data (Report No. 15-UM-036). National Surface Transportation Safety Center for Excellence, Virginia Tech Transportation Institute, Blacksburg, VA
16. Guo F, Fang Y, Antin JF (2015b) Older driver fitness-to-drive evaluation using naturalistic driving data. J Saf Res 54:49–54. https://doi.org/10.1016/j.jsr.2015.06.013
17. Hatakka M, Keskinen E, Gregersen NP, Glad A, Hernetkoski K (2002) From control of the vehicle to personal self-control: broadening the perspectives to driver education. Transp Res Part F: Traffic Psychol Behav 5:201–215. https://doi.org/10.1016/S1369-8478(02)00018-9
18. Horrey WJ, Lesch MF, Mitsopoulos-Rubens E, Lee JD (2015) Calibration of skill and judgment in driving: development of a conceptual framework and the implications for road safety. Accid Anal

Prev 76:25–33. https://doi.org/10.1016/j.aap.2014.12.017

19. Horswill MS, Sullivan K, Lurie-Beck JK, Smith S (2013) How realistic are older drivers' ratings of their driving ability? Accid Anal Prev 50:130–137. https://doi.org/10.1016/j.aap.2012.04.001
20. Ichikawa M, Inada H, Nakahara S (2020) Increased traffic injuries among older unprotected road users following the introduction of an age-based cognitive test to the driver's license renewal procedure in Japan. Accid Anal Prev 136:105440. https://doi.org/10.1016/j.aap.2020.105440
21. Inada H, Tomio J, Nakahara S, Ichikawa M (2023) Association between mandatory cognitive testing for license renewal and motor vehicle collisions and road injuries among older drivers in Japan. J Am Geriatr Soc 71:1145–1155. https://doi.org/10.1111/jgs.18157
22. Jessen F, Amariglio RE, van Boxtel M, Breier M, Cainelli E, Dufouil C, Ellis KA, van der Flier WM, Glodzik L, van Harten AC, de Leon MJ, McHugh P, Mielke MM, Molinuevo JL, Mosconi L, Osorio RS, Perrotin A, Petersen RC, Rabin LA, Rami L, Reisberg B, Rentz DM, Sachdev PS, de la Sayette V, Saykin AJ, Scheltens P, Shulman MB, Slavin MJ, Sperling RA, Stewart R, Uspenskaya O, Vellas B, Visser PJ, Wagner M Subjective Cognitive Decline Initiative (SCD-I) Working Group, 2014. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimer's Dement 10, 844–852. https://doi.org/10.1016/j.jalz.2014.01.001
23. Keay L, Munoz B, Duncan DD, Hahn D, Baldwin K, Turano KA, Munro CA, Bandeen-Roche K, West SK (2013) Older drivers and rapid deceleration events: Salisbury Eye Evaluation Driving Study. Accid Anal Prev 58:279–285. https://doi.org/10.1016/j.aap.2012.06.002
24. Kurita S, Doi T, Harada K, Morikawa M, Nishijima C, Fujii K, Kakita D, Shimada H (2024) Subjective memory concerns and car collisions: a cross-sectional cohort study among older Japanese drivers. Heliyon 10:e33080. https://doi.org/10.1016/j.heliyon.2024.e33080
25. Lakens D (2017) Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Social Psychol Personality Sci 8:355–362. https://doi.org/10.1177/1948550617697177
26. Lakens D, Scheel AM, Isager PM (2018) Equivalence testing for psychological research: a tutorial. Adv Methods Practices Psychol Sci 1:259–269. https://doi.org/10.1177/2515245918770963
27. Li G, Braver ER, Chen L-H (2003) Fragility versus excessive crash involvement as determinants of high death rates per vehicle-mile of travel among older drivers. Accid Anal Prev 35:227–235. https://doi.org/10.1016/S0001-4575(01)00107-5
28. Li G, Eby DW, Santos R, Mielenz TJ, Molnar LJ, Strogatz D, Betz ME, DiGuiseppi C, Ryan LH, Jones V, Pitts SI, Hill LL, DiMaggio CJ, LeBlanc D, Andrews HF (2017) LongROAD Research Team, Longitudinal Research on Aging Drivers (LongROAD): study design and methods. Injury Epidemiology 4, 22. https://doi.org/10.1186/s40621-017-0121-z
29. Liu D, Chihuri S, Andrews HF, Betz ME, DiGuiseppi C, Eby DW, Hill LL, Jones V, Mielenz TJ, Molnar LJ, Strogatz D, Lang BH, Li G (2024) Diabetes mellitus and hard braking events in older adult drivers. Injury Epidemiol 11:22. https://doi.org/10.1186/s40621-024-00508-2
30. Marottoli RA, Richardson ED (1998) Confidence in, and self-rating of, driving ability among older drivers. Accid Anal Prev 30:331–336. https://doi.org/10.1016/S0001-4575(97)00100-0

31. Mathias JL, Lucas LK (2009) Cognitive predictors of unsafe driving in older drivers: a meta-analysis. Int Psychogeriatr 21:637–653. https://doi.org/10.1017/S1041610209009119
32. Merickel J, High R, Dawson J, Rizzo M (2019) Real-world risk exposure in older drivers with cognitive and visual dysfunction. Traffic Inj Prev 20:S110–S115. https://doi.org/10.1080/15389588.2019.1688794
33. Michon JA (1985) A critical view of driver behavior models: what do we know, what should we do? In: Evans L, Schwing RC (eds) Human Behavior and Traffic Safety. Plenum, New York, pp 485–524. https://doi.org/10.1007/978-1-4613-2173-6_19
34. Molnar LJ, Eby DW (2008) The relationship between self-regulation and driving-related abilities in older drivers: an exploratory study. Traffic Inj Prev 9:314–319. https://doi.org/10.1080/15389580801895319
35. Reger MA, Welsh RK, Watson GS, Cholerton B, Baker LD, Craft S (2004) The relationship between neuropsychological functioning and driving ability in dementia: a meta-analysis. Neuropsychology 18:85–93. https://doi.org/10.1037/0894-4105.18.1.85
36. Ross LA, Clay OJ, Edwards JD, Ball KK, Wadley VG, Vance DE, Cissell GM, Roenker DL, Joyce JJ (2009) Do older drivers at-risk for crashes modify their driving over time? Journals Gerontology: Ser B 64:163–170. https://doi.org/10.1093/geronb/gbn034
37. Seelye A, Mattek N, Sharma N, Witter P, Brenner A, Wild K, Dodge H, Kaye J (2017) Passive assessment of routine driving with unobtrusive sensors: a new approach for identifying and monitoring functional level in normal aging and mild cognitive impairment. J Alzheimer's Disease 59:1427–1437. https://doi.org/10.3233/JAD-170116
38. Stefanidis KB, Mieran T, Schiemer C, Freeman J, Truelove V, Summers MJ (2023) Cognitive correlates of reduced driving performance in healthy older adults: a meta-analytic review. Accid Anal Prev 193:107337. https://doi.org/10.1016/j.aap.2023.107337
39. Sugishita M, Koshizuka Y, Sudou S, Sugishita K, Hemmi I, Karasawa H, Ihara M, Asada T, Mihara B (2018) The validity and reliability of the Japanese version of the Mini-Mental State Examination (MMSE-J) with the original procedure of the attention and calculation task (2001). Japanese J Cogn Neurosci 20:91–110 (in Japanese). https://doi.org/10.11253/ninchishinkeikagaku.20.91
40. Tsoi KKF, Chan JYC, Hirai HW, Wong SYS, Kwok TCY (2015) Cognitive tests to detect dementia: a systematic review and meta-analysis. JAMA Intern Med 175:1450–1458. https://doi.org/10.1001/jamainternmed.2015.2152

Figures

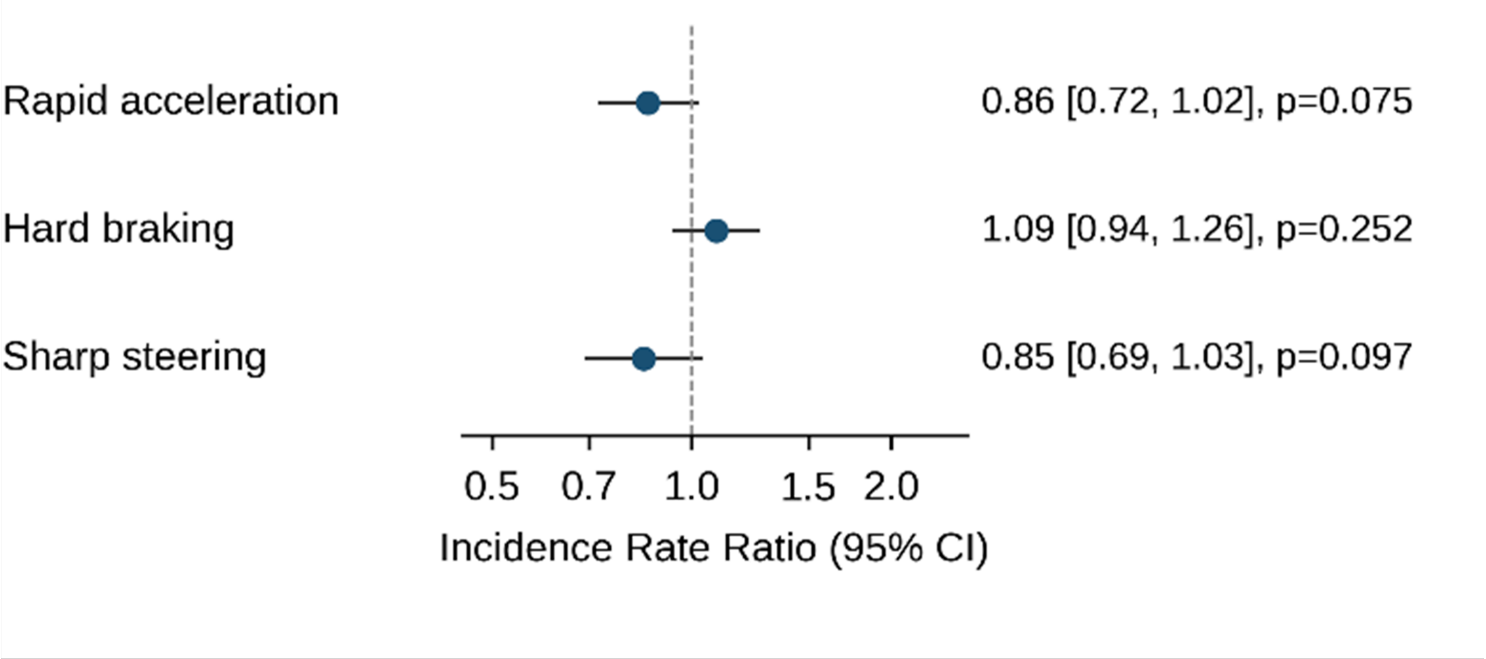


Figure 1

Association between MMSE score and the rate of hazardous driving events. Incidence rate ratios (IRRs) with 95% confidence intervals (CIs) are shown for the three event types (rapid acceleration, hard braking, sharp steering) from separate negative binomial regression models (n = 192). Each model estimated the IRR per one-point increase in MMSE score, adjusted for age, sex (male as the reference), and years of education, with the natural logarithm of driving time (hours) entered as an offset. The dashed vertical line marks the null value (IRR = 1). The x-axis is displayed on a logarithmic scale. Exact IRRs, 95% CIs, and p-values are reported in Table 2.

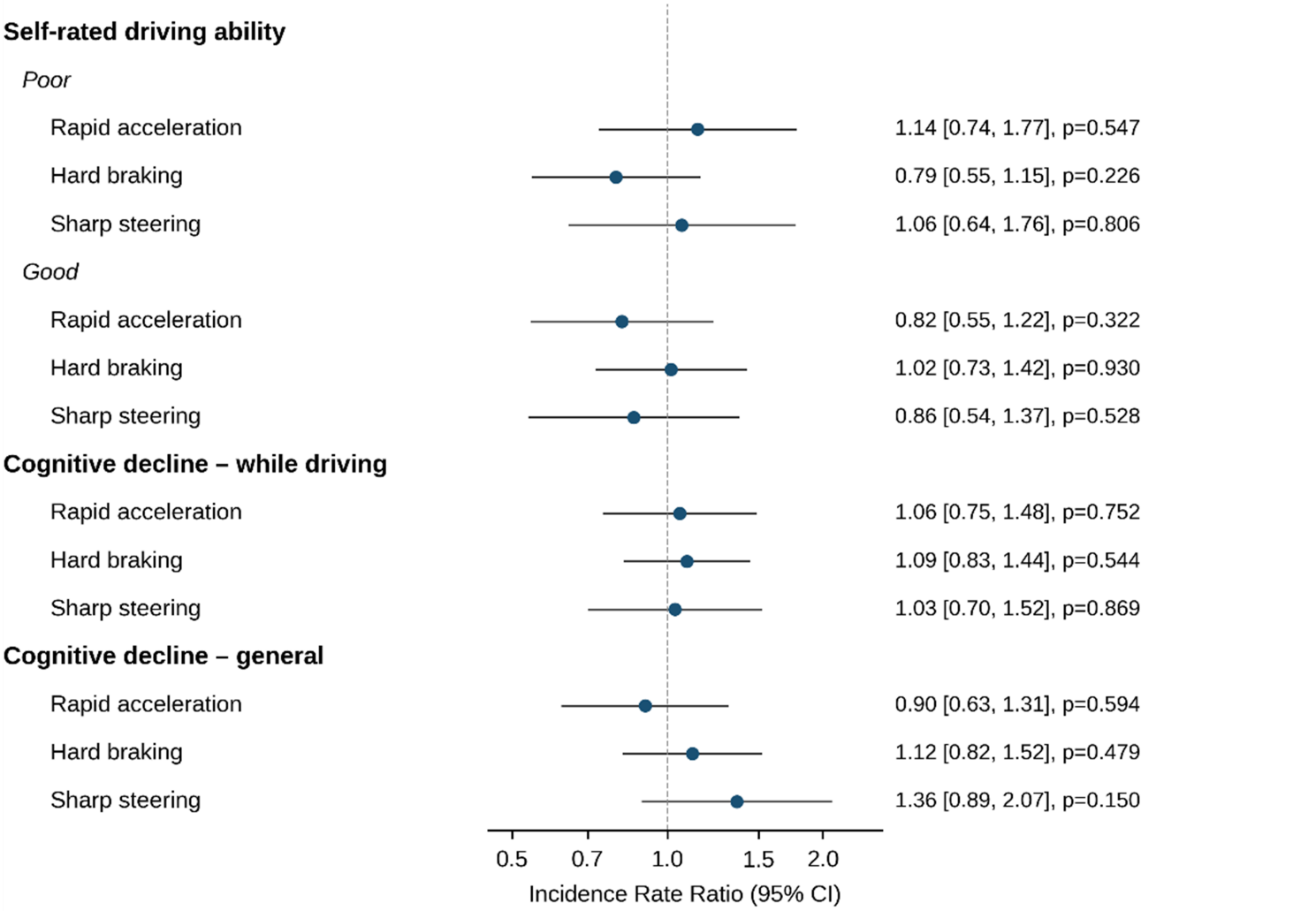


Figure 2

Associations between self-perceived cognitive decline and the rate of hazardous driving events. Incidence rate ratios (IRRs) with 95% confidence intervals (CIs) are shown for two binary self-appraisal indicators—self-perceived cognitive decline while driving (yes versus no) and self-perceived cognitive decline in general (yes versus no)—each against the three event types (rapid acceleration, hard braking, sharp steering), from negative binomial regression models (n = 192). Self-rated driving ability (poor / neither / good) was evaluated with a single joint Wald test (2 df); because a unique IRR is not defined for a three-level categorical predictor, it is not plotted and is reported in Table 2. All models adjusted for age, sex (male as the reference), and years of education, with the natural logarithm of driving time (hours) entered as an offset. The dashed vertical line marks the null value (IRR = 1). The x-axis is displayed on a logarithmic scale.