Supplementary Material

Care-Management Proximity and the Survival of Home-Visit Nursing Agencies in Japan, 2020–2025: A National Cohort Study

Hiroyuki Tajima, Faculty of Nursing, Shumei University, Japan

All values in this Supplementary Material are computed from the analytic dataset used for the main text, in which corporate identity is assigned by the linkage described in Supplementary Methods S1. Supplementary Table S19 compares the main estimates with those obtained without the linkage. Every accelerated failure time model uses standard errors clustered on the operating corporation, as in the main text. The analysis code and the derived dataset are publicly available at Zenodo (https://doi.org/10.5281/zenodo.21357827; Supplementary Note S9).

Abbreviations: AFT, accelerated failure time; AIC, Akaike information criterion; CI, confidence interval; DOI, digital object identifier; HR, hazard ratio; LTCI, long-term care insurance; NPO, non-profit organisation; SD, standard deviation; TR, time ratio; URL, uniform resource locator.

Supplementary Methods S1: Data fields, provider form and corporate identifier

Data fields. Each record in the open-data files carries 24 fields: a prefecture or municipality code, a row number, prefecture, municipality, facility name and name in kana, service type, address and building name, latitude and longitude, telephone and fax numbers, the operating corporation's national corporate number and registered name, a unique 10-digit insurance provider number, days of operation and notes on them, capacity, a URL, a flag for services that older people and people with disabilities can use together (共生型), two fields on designation under long-term care insurance and under disability welfare services, and a free-text remarks field. There is no measure of staffing, caseload or client volume, and no designation date or date of commencement.

Provider form. The third digit of the provider number distinguishes freestanding stations (5, 6, 7) from hospitals and clinics under deemed designation (1, 3). Across all facilities ever recorded under service 130 (n=22,350), the rule classifies 20,757 as stations and 1,593 as medical institutions. Validation against facility names supports it: 94.0% of station-coded facilities carry a nursing-station term (訪問看護, ナースステーション) and 2.7% a hospital or clinic term, while 90.0% of medical-institution-coded facilities carry a hospital or clinic term (病院, 医院, クリニック, 診療所).

Corporate identifier. The 13-digit national corporate number (法人番号) is reported on some, but not all, facility records, and the records of one corporation can differ in whether they report it. In the December 2020 files of all 35 service types (222,764 records), 124,457 records reported a number. Every record was assigned a corporate identifier in four steps, using only these baseline records so that no information from later waves enters the identifier:

1. *Direct*: the number reported on the record.
2. *Linked*: for a record without a number, the number corresponding to its prefecture and registered corporate name, provided that this combination corresponded to exactly one number among the baseline records that reported one. Names were compared after Unicode compatibility

normalisation (NFKC) and removal of all white space, so that variants differing only in character width or spacing match.

3. *Proxy*: otherwise, the combination of prefecture and the registered name as recorded.
4. *Standalone*: Where neither a number nor a name was reported, the record was treated as a single-office corporation.

Of the 97,988 baseline records without a number but with a name, 33,119 were linked to a number in step 2, 2,025 were not linked because their prefecture and name corresponded to more than one number, and 62,844 had no counterpart among the records reporting a number. Service scope, corporate scale, care-management ownership and the distances to the corporation's offices were then computed over all records sharing an identifier. The submitted version of the manuscript used steps 1, 3 and 4 only, so that records of one corporation that differed in whether they reported the number were treated as different corporations; the consequences of that difference are reported in Supplementary Table S19.

Table S1. Corporate-identifier type, for the baseline cohort and the analytic cohort of freestanding stations.

Identifier	Baseline cohort (n=13,032)	%	Analytic cohort (n=11,778)	%
Direct (national corporate number reported)	7,388	56.7	6,949	59.0
Linked (number assigned from a prefecture and name corresponding to one number)	1,701	13.1	1,544	13.1
Proxy (prefecture and registered name)	3,924	30.1	3,271	27.8
Standalone (neither number nor name reported)	19	0.1	14	0.1

In the analytic cohort, 8,493 stations (72.1%) belong to a corporation identified by a number, reported or linked; these are the stations used for the time-varying analysis (Supplementary Note S20). Corporate identity rests on the registered name, either linked to a number or alone, for 40.9% of the analytic cohort.

The 14 standalone stations report neither a corporate number nor a corporate name. They are treated as single-office corporations, which assigns each of them a service scope of one, a corporate scale of one and no care-management office (§2.4 of the main text). This is a coding of missing information rather than an observation. Removing these stations changes no time ratio of the main model by more than 0.001. None of them enters the analyses restricted to stations whose corporation ran a care-management office; in the three-group model of the main text (Table 4) they are among the non-owners.

Validation of the linkage key. Records that report a number have a known corporate identity, so they show how often a prefecture and name identify a single corporation. Among the 124,450 baseline records reporting both a number and a name, the 46,136 distinct combinations of prefecture and normalised name included 45,413 (98.43%) that corresponded to exactly one number; 723 (1.57%) corresponded to between two and six numbers and covered 6,281 records (5.05%). Without normalisation there are 55,831 combinations, of which 98.88% correspond to one number (3.45% of records affected), because variants of one name that differ only in spacing then count as different combinations. Ambiguous combinations are dominated by national chains operating several separately incorporated entities under one trade name and by short generic names; they were not used for linking.

Two kinds of error remain. *Over-merging* assigns two corporations that share a prefecture and name to one identifier. It inflates service scope and corporate scale and can attribute another corporation's care-management office to a station. For linked and proxy records it cannot be checked directly, because the corporation behind a record without a number is unknown; the validation above shows that names shared by more than one numbered corporation are uncommon and were excluded from the linkage. *Over-*

splitting assigns one corporation to more than one identifier: a proxy record whose registered name is written differently from the corporation's other records, or a corporation operating in several prefectures without reporting its number. It deflates scope and scale and can hide a care-management office of the same corporation. If the hidden office is near the station, the station is classified as further away or as a non-owner, which works against the proximity hypothesis; if the hidden office is the corporation's only, distant office, an owner further away is classified as a non-owner, which makes the two groups look alike and so favours the reading that ownership without proximity confers no advantage. Records that carry the number, reported or linked, are joined across prefectures, because the number is national. A record that neither reports the number nor can be linked within its own prefecture remains a separate identifier, so a corporation is still split where none of its records in a prefecture reports the number, or where its name is written differently on different records. The analyses restricted to stations with a directly reported number (Supplementary Table S11) examine the consequences.

Supplementary Methods S2: Validation of the agency-vintage proxy

The disclosure files record no designation date or date of commencement of operations (§2.1 of the main text), so agency age cannot be taken from them. Provider numbers are issued sequentially within jurisdiction blocks, and the manuscript uses a rank derived from them as a proxy: Stations are grouped by the first six digits of the 10-digit provider number, ranked by full number within that block, and rescaled to the unit interval, using only providers present in December 2020 so that no information from later waves enters the covariate.

Validation against published commencement dates within the cohort. Each facility's page in the Long-Term Care Service Information Disclosure System publishes its date of commencement of operations (事業開始年月日), although the bulk open-data files do not. A random sample of 800 of the 11,778 stations in the analytic cohort was drawn with a fixed seed, and the published date of commencement was retrieved from each station's page on 21 September 2026. Of the 800, 602 had a published date, of which only two had closed by December 2025; of the 198 without a published page, 175 had closed under the wider definition and 144 under the primary definition (Supplementary Note S16). Among the 602, the rank correlated strongly with the date of commencement (Spearman's ρ=0.78; Kendall's τ=0.60), and median commencement rose steadily across quintiles of the rank (Table S2, first section). Three stations recorded a commencement date after December 2020 and were retained.

A second test among later entrants. The 9,318 agencies that entered the registry after December 2020 lie outside the analytic cohort, and for them the wave of first appearance approximates the date of entry. Applying the same ranking procedure to these agencies and comparing the rank against entry wave (Table S2, second section) gives Spearman's ρ=0.79 and Kendall's τ=0.65, and median rank rises steadily across entry waves, from 0.10 in June 2021 to 0.83 in December 2025.

Table S2. Validation of the provider-number rank.

Comparison	Group	Median	Interquartile range	n
Cohort stations: published date of commencement, by quintile of the rank (n=602)				
	Quintile 1 (lowest rank)	March 2000	December 1996 to January 2007	121
	Quintile 2	October 2009	October 2001 to March 2014	121

Comparison	Group	Median	Interquartile range	n
	Quintile 3	October 2014	January 2012 to April 2016	120
	Quintile 4	February 2017	April 2015 to May 2018	120
	Quintile 5 (highest rank)	July 2019	November 2017 to April 2020	120
Agencies entering after baseline: rank within block, by entry wave (n=9,318)				
	June 2021	0.099	0.050–0.167	574
	December 2021	0.167	0.096–0.250	944
	June 2022	0.250	0.156–0.344	695
	December 2022	0.339	0.250–0.438	984
	December 2023	0.477	0.350–0.562	1,865
	December 2024	0.625	0.500–0.731	1,587
	December 2025	0.833	0.703–0.917	2,669

The two results are consistent. Within the cohort, stations founded in the 1990s and those founded in the late 2010s occupy clearly different parts of their numbering blocks, and the rank separates them well. Among agencies entering within the same five-year window, the rank also orders the wave of entry. Because numbers are reissued and blocks fill at different rates in different jurisdictions, the rank is an ordinal proxy for agency age rather than a date. Because the validation sample consists of stations still listed in 2026, it cannot show whether the rank orders closed stations equally well.

The proxy is used as an adjustment covariate in every model, and all primary models are reported with and without it (§3.7 of the main text; Supplementary Note S13). Omitting it moves every organisational and proximity estimate further from the null, so its inclusion is the conservative choice. Its own association with closure is discussed in §4.1 of the main text.

Supplementary Table S3. Full-cohort sensitivity analysis (provider form controlled directly)

Restoring the 1,254 deemed-designated medical institutions excluded from the main analytic cohort (§2.2 of the main text) and controlling for provider form directly:

Covariate	Time ratio	95% CI	p
Owns a care-management office	1.365	1.168–1.596	<0.001
Service scope (per log service type)	1.232	1.065–1.425	0.005
Corporate scale (per log office)	0.869	0.789–0.956	0.004
Agency-vintage proxy (highest vs lowest rank)	0.552	0.464–0.657	<0.001
Deemed-designated medical institution (vs station)	0.302	0.269–0.340	<0.001

n=13,032; 2,684 closures. Weibull accelerated failure time model with prefecture fixed effects, unpenalised, with closure defined as in §2.3 of the main text. Exits of the 1,254 medical institutions were not classified (Supplementary Table S10), so every disappearance of their provider numbers is counted as

a closure. The organisational associations keep their direction and significance; deemed designation is itself the strongest single predictor in the pooled cohort, which is why the two provider forms are analysed separately.

Supplementary Table S4. Multi-service-corporation sensitivity analysis

Restricting to stations whose operating corporation ran two or more LTCI service types in December 2020, to examine whether single-service corporations, which take the lowest value of both service scope and corporate scale, drive those associations:

Covariate	Time ratio	95% CI	p
Owns a care-management office	1.445	1.210–1.726	<0.001
Service scope (per log service type)	1.060	0.829–1.355	0.642
Corporate scale (per log office)	0.984	0.813–1.192	0.872
Agency-vintage proxy (highest vs lowest rank)	0.403	0.314–0.517	<0.001

n=7,793; 1,240 closures. The care-management ownership association is unchanged. Within this restricted sample neither service scope nor corporate scale is associated with closure, so in the full cohort both associations rest largely on the contrast between single-service and multi-service corporations (§3.7 and §4.1 of the main text).

Supplementary Table S5. Joint versus separate placebo models (distance to own care management vs own day care)

The two distance measures are correlated (Spearman's ρ=0.262, n=2,490), so both specifications are reported.

Model	Exposure (per log-km)	Time ratio	95% CI	p
Separate models	Distance to own care management	0.940	0.844–1.047	0.260
Separate models	Distance to own day care	1.089	0.979–1.213	0.118
Joint model	Distance to own care management	0.887	0.796–0.989	0.030
Joint model	Distance to own day care	1.128	1.009–1.261	0.034

n=2,490 for all four rows (stations whose corporation ran both a care-management office and a day-care centre, with usable coordinates for each). A time ratio below 1 means that greater distance shortens expected survival. Entered separately, neither distance is associated with closure in this subsample. Entered together, the two separate in opposite directions: Greater distance from the care-management office is associated with shorter survival and greater distance from the day-care centre with longer survival. Under the wider closure definition, the separate models give 0.880 (95% CI 0.789–0.982) for care management and 1.078 (95% CI 0.974–1.193) for day care. The day-care comparison therefore does not show a general advantage of being near the corporation's other services, but in this subsample the care-management gradient appears under the primary definition only when distance to the day-care centre is held constant. The placebo used in the main text is whether a care-management office run by another corporation lies within 0.5 km (Table 4 of the main text; Supplementary Table S15).

Supplementary Table S6. Legal form: associations with closure before and after adjustment (reference category: joint-stock company)

Legal form	n	Care-management ownership, %	TR, prefecture only	p	Adjusted TR*	95% CI	p
Medical corporation	2,568	75.3	1.526	<0.001	0.970	0.819–1.149	0.727
Other or unclassified	945	41.3	1.387	0.004	1.227	0.993–1.516	0.058
Limited company (有限会社)	750	52.9	0.844	0.110	0.667	0.550–0.810	<0.001
Social-welfare corporation	661	89.0	1.045	0.756	0.640	0.480–0.853	0.002
Limited-liability company (合同会社)	631	16.5	0.948	0.649	1.053	0.846–1.311	0.645
Incorporated association	623	68.9	1.489	0.007	1.098	0.815–1.479	0.538
Cooperative	272	96.0	1.377	0.178	0.747	0.465–1.200	0.229
Incorporated foundation	172	77.9	1.996	0.025	1.266	0.692–2.317	0.444
NPO corporation	133	42.1	0.853	0.476	0.714	0.461–1.107	0.132

The first time ratio is from a model containing legal form and prefecture fixed effects only. * Adjusted additionally for care-management ownership, service scope, corporate scale and the agency-vintage proxy. Joint-stock company (n=5,007; 39.8% care-management ownership) is the reference category throughout; n=11,778 stations and 2,141 closures, of which the 16 stations with no recorded corporate name form a separate category that is not shown (two of them report a corporate number, which is why they outnumber the 14 standalone stations of Supplementary Methods S1, which report neither). Closure is defined as in §2.3 of the main text. Before adjustment, medical corporations, incorporated associations and incorporated foundations appear to outlast joint-stock companies. After adjustment no non-profit form retains a survival advantage over joint-stock companies; the social-welfare corporation estimate reverses direction (0.640, 95% CI 0.480–0.853), and limited companies, already below 1 before adjustment, are associated with significantly shorter survival after it (0.667, 95% CI 0.550–0.810). Care-management ownership varies widely across legal forms (from 16.5% to 96.0%), which is why the unadjusted comparison confounds legal form with the service portfolio. The labels of two company forms have been corrected from the submitted version, in which the 有限会社 row was labelled ‘limited-liability company’ and the 合同会社 row ‘limited/joint-stock partnership’.

Supplementary Table S7. Proximity and ownership with corporate archetype additionally controlled

Model	Exposure	Time ratio	95% CI	p
Owners, without archetype control	Within walking distance (<0.5	1.449	1.218–	<0.001

Model	Exposure	Time ratio	95% CI	p
	km)		1.725	
Owners, with archetype control	Within walking distance (<0.5 km)	1.417	1.190–1.686	<0.001
All stations, without archetype control	Owns a care-management office	1.462	1.217–1.756	<0.001
All stations, with archetype control	Owns a care-management office	1.533	1.279–1.837	<0.001
All stations, with archetype control	Health-facility-affiliated archetype	1.290	0.975–1.707	0.074
All stations, with archetype control	Nursing-home-affiliated archetype	0.789	0.577–1.078	0.136
All stations, with archetype control	Rehabilitation-centred archetype	1.474	1.124–1.933	0.005
All stations, with archetype control	Single-service corporation	1.089	0.869–1.364	0.459

Owners: n=6,264 stations whose corporation ran a care-management office with usable coordinates (916 closures). All stations: n=11,778 (2,141 closures). Reference archetype: home-care coordination, the modal archetype; corporations running a single service type form their own category. Models are adjusted for service scope, corporate scale, the vintage proxy and prefecture fixed effects. Controlling for a corporation's broader service-portfolio archetype does not materially change either the proximity or the ownership association. Service scope, which the archetypes partly encode, is not significant once they are controlled (1.184, 95% CI 0.913–1.537).

Supplementary Note S8: Agency age as an alternative explanation of the proximity association

An alternative account of the proximity finding is confounding by agency age: Stations near their corporation's care-management office might be older, more established branches, and age rather than proximity might protect them. Two pieces of evidence weigh against this as a sole explanation. First, distance to the corporation's own care-management office is only weakly correlated with the vintage proxy (Spearman's ρ=0.183, n=6,264), which tracks published commencement dates (Supplementary Methods S2). Second, every distance model in the main text holds the proxy constant, and adding it attenuates the proximity estimates only modestly (within walking distance, time ratio 1.589 without the proxy and 1.449 with it; within-corporation comparison, hazard ratio 0.631 without and 0.665 with it, the latter in Table 5 of the main text). Agency age therefore accounts for part of the unadjusted proximity contrast but not for its persistence after adjustment, although the validation of the proxy covers surviving stations only.

Supplementary Note S9: Reproducibility

The analysis pipeline (data download, corporate-identity linkage, panel construction, classification of exits, survival-dataset assembly, community detection, analysis, simulation, figures and tables) is scripted, resumable and parallelised, and an automated audit checks the reported estimates against the analysis output. The pipeline, the audit and the derived analytic dataset are publicly available at Zenodo (https://doi.org/10.5281/zenodo.21357827). The facility-page records retrieved in September 2026 for the validation of the vintage proxy (Supplementary Methods S2) and the check of closures (Supplementary

Note S16) were not retained, so those counts cannot be recomputed from the deposit. The few values that the deposited code reproduces only approximately are listed, with the size of each difference, in the deposit's README.

Supplementary Table S10. Classification of exits, and estimates under alternative closure definitions

A disappearing provider number need not mean that a locality lost nursing: A station may be renumbered, its designation transferred (事業承継), its site taken over by another provider, or its clients absorbed by a nearby branch of the same corporation. Using the full eight-wave panel (all 22,350 provider numbers, with coordinates), each of the 2,543 station exits with valid coordinates was classified as: (i) a new provider number at the same site, meaning a home-visit nursing provider number of any corporation, first recorded after December 2020, within 100 m of the closed station and first appearing at or after the last wave at which the station was present; (ii) a home-visit nursing provider of the same corporation (identified as in Supplementary Methods S1) operating within 500 m in December 2025; or (iii) closure (neither). The five exits without valid coordinates could not be classified and are counted as closures.

Closure type	n	% of 2,548 exits
New provider number at the same site (≤100 m)	378	14.8
Home-visit nursing provider of the same corporation operating nearby (≤500 m)	70	2.7
Either of the above	407	16.0
Closure (primary outcome): neither of the above	**2,141**	**84.0**

Categories (i) and (ii) overlap, because 41 exits met both criteria; the union (407) is therefore smaller than the sum of the two categories. Exits of the 1,254 deemed-designated medical institutions, which lie outside the analytic cohort, were not classified.

The main text uses the primary definition throughout, because an exit that leaves a new provider number at the same site, or a provider of the same corporation nearby, is often an administrative event rather than the end of a station and in any case leaves provision at or near the site; where such an exit was in fact the end of a station, treating it as censored makes the proximity estimates smaller, as shown below. A closure under this definition is the loss of a station, not necessarily a loss of local capacity, which the main text examines at municipal level (§3.2). The table below reports both definitions side by side so that the consequence of the choice is visible; all estimates are unpenalised.

Estimate	Primary definition	Wider definition (all disappearances)
Within walking distance (Weibull AFT, prefecture fixed effects, n=6,264)	TR 1.449 (95% CI 1.218–1.725, p<0.001)	TR 1.610 (95% CI 1.352–1.918, p<0.001)
Within-corporation comparison (stratified Cox, n=2,393 in 650 corporations)	HR 0.665 (95% CI 0.500–0.883, p=0.005)	HR 0.570 (95% CI 0.441–0.738, p<0.001)
Continuous log-distance (Weibull AFT, prefecture fixed effects, n=6,264)	TR 0.903 (95% CI 0.843–0.967, p=0.004)	TR 0.866 (95% CI 0.812–0.924, p<0.001)

The proximity estimates are larger under the wider definition because exits of stations further from their corporation's care-management office were more often followed by a new home-visit nursing provider number at the same site: 21.2% of their 415 exits, against 13.3% of the 701 exits within walking distance.

A provider of the same corporation nearby was similarly uncommon in the two groups (3.9% and 4.1%). Counting these exits as closures therefore adds proportionally more events among stations further away, and the primary definition is the conservative one for this paper's argument. Under it, the walking-distance contrast, the within-corporation comparison and the continuous distance gradient all remain significant.

The organisational associations are barely affected by the choice of definition (care-management ownership 1.462 under the primary definition against 1.429 under the wider one; service scope 1.235 against 1.217; corporate scale 0.870 against 0.874).

The new provider number at the same site. Rule (i) counts a new number of any corporation, first appearing at or after the last wave at which the station was present with no limit on how much later, so it covers replacement at the same site, possibly by another operator, as well as renumbering and succession. The table gives, for the exits it reclassified, when the new number first appeared.

First appearance of the new number	Exits	Cumulative
At the last wave at which the closed station was present	38	38
At the wave at which the station disappeared	149	187
One wave after the station disappeared	102	289
Two to six waves after the station disappeared	89	378

The new number carried the same corporate identifier as the closed station for 42 of the 378 exits. For the others the identifier differed, which may reflect a transfer of the business to another corporation, a change of operator at the site, or a difference in how the two records identify the corporation; the data cannot tell these apart. Waves are half a year apart to December 2022 and a year apart thereafter.

Limiting rule (i) by timing. To see how much the estimates depend on new numbers that appeared well after the exit, rule (i) was limited to new numbers that had appeared by the wave at which the station disappeared, or by one wave after it; rule (ii) was kept.

Estimate	Rule (i) as defined (primary)	By the wave of exit	By one wave after exit
Closures	2,141	2,303	2,208
Owns a care-management office, TR	1.462 (1.217–1.756)	1.483 (1.226–1.795)	1.533 (1.271–1.850)
Service scope (per log service type), TR	1.235 (1.048–1.456)	1.237 (1.049–1.458)	1.226 (1.036–1.449)
Corporate scale (per log office), TR	0.870 (0.785–0.964)	0.867 (0.785–0.959)	0.867 (0.781–0.963)
Within walking distance (owners), TR	1.449 (1.218–1.725)	1.555 (1.296–1.865)	1.467 (1.230–1.749)
Continuous log-distance (owners), TR	0.903 (0.843–0.967)	0.886 (0.827–0.949)	0.901 (0.841–0.966)
Own office ≥0.5 km vs none, TR	1.144 (0.907–1.443)	1.109 (0.868–1.416)	1.189 (0.937–1.509)
Own office <0.5 km vs none, TR	1.632 (1.352–1.969)	1.694 (1.396–2.055)	1.719 (1.418–2.084)
Within-corporation comparison,	0.665 (0.500–0.883)	0.650 (0.496–	0.667 (0.503–

Estimate	Rule (i) as defined (primary)	By the wave of exit	By one wave after exit
HR		0.850)	0.882)

Models as in the main text (Weibull accelerated failure time models with prefecture fixed effects and corporation-clustered standard errors; the within-corporation comparison is a Cox model stratified on the corporation, with model-based standard errors). The walking-distance estimate under either limit lies between those under the primary and the wider definitions, every proximity estimate remains significant, and owners further from their office do not differ significantly from non-owners under any of the three rules.

Supplementary Table S11. Robustness to the corporate identifier: stations with a directly reported corporate number

Because corporate identity rests on the registered name for 40.9% of stations in the analytic cohort (Table S1), every estimate was repeated on the stations whose own record reported a national corporate number. This analysis was also reported in the submitted version, with the identifier used there.

Covariate	Full cohort TR	95% CI	Direct-number TR	95% CI
Owns a care-management office	1.462	1.217–1.756	1.595	1.251–2.033
Service scope (per log service type)	1.235	1.048–1.456	1.178	0.966–1.438
Corporate scale (per log office)	0.870	0.785–0.964	0.854	0.765–0.953
Agency-vintage proxy (highest vs lowest rank)	0.505	0.415–0.613	0.457	0.352–0.593
Within walking distance (<0.5 km)	1.449	1.218–1.725	1.508	1.217–1.869

Full cohort n=11,778 (n=6,264 for the distance model); direct-number subsample n=6,949 (1,183 closures; n=4,110 and 564 closures for the distance model). The ownership and proximity estimates are larger in the subsample than in the full cohort, and corporate scale and the vintage proxy remain significant, but service scope does not (1.178, 95% CI 0.966–1.438, p=0.106).

Ownership and proximity together, by identifier. The three-group model of the main text (Table 4) was re-estimated on the stations with a directly reported number and on those whose corporation was identified by a number, reported or linked. The comparison bears on the over-splitting described in Supplementary Methods S1, which would make owners further away look like non-owners.

Sample	n (closures)	Own office ≥0.5 km vs none	Own office <0.5 km vs none	<0.5 km vs ≥0.5 km
All stations (main analysis)	11,757 (2,135)	1.144 (0.907–1.443)	1.632 (1.352–1.969)	1.426 (1.194–1.702)
Corporation identified by a number, reported or linked	8,475 (1,470)	1.152 (0.880–1.507)	1.673 (1.339–2.091)	1.453 (1.201–1.758)
Directly reported number	6,938 (1,179)	1.280 (0.944–1.736)	1.805 (1.409–2.311)	1.409 (1.123–1.769)

Time ratios with 95% confidence intervals from Weibull accelerated failure time models with prefecture fixed effects, adjusted for service scope, corporate scale and the vintage proxy; stations whose corporation ran a care-management office without usable coordinates are excluded. The advantage of owners within walking distance over owners further away is similar in all three samples (1.409 to 1.453). Owners further away do not differ significantly from non-owners in any of the three samples, although among stations with a directly reported number the point estimate is higher (1.280, p=0.112). The absence of an advantage beyond walking distance is therefore less secure than the difference between near and distant owners. In this model the near-distant contrast among stations with a directly reported number (1.409) is smaller than in the owners-only model above (1.508), because scope, scale and the vintage proxy are estimated on all stations rather than on owners only.

Supplementary Note S12: Organisational archetypes

To characterise the broader service ecosystem, and as a robustness check on the proximity association reported in §3.5 of the main text, corporations were grouped into archetypes by service-portfolio similarity, using the December 2020 records of all 35 service types and the corporate identifier of Supplementary Methods S1. Corporations with an identical set of LTCI service types formed a single node (4,178 distinct portfolios among 35,794 corporations running two or more service types); portfolios connected by a Jaccard-similarity edge of at least 0.5 were partitioned into communities by the Louvain method (Blondel *et al.*, 2008) with a fixed random seed, giving nine communities with modularity 0.482. Four archetypes (home-care coordination, health-facility-affiliated, rehabilitation-centred and nursing-home-affiliated) accounted for 99.98% of the multi-service corporations; corporations running a single service type form a fifth category. Controlling for archetype left the proximity and ownership associations essentially unchanged (Table S7; Table 4 of the main text).

Supplementary Note S13: Sensitivity analyses (full results)

Wider closure definition. Counting every disappearance of a provider number, including exits with a new provider number at the same site or a provider of the same corporation nearby, the organisational associations are almost unchanged (care-management ownership TR=1.429, service scope 1.217, corporate scale 0.874), while the proximity associations are larger (within walking distance TR=1.610; within-corporation HR=0.570; continuous log-distance TR=0.866; Supplementary Table S10). Limiting the same-site rule to new numbers that had appeared by the wave of exit gives intermediate estimates (within walking distance TR=1.555; Supplementary Table S10).

Without the vintage proxy. Omitting the proxy strengthened every estimate (care-management ownership TR=1.579, 95% CI 1.314–1.898; service scope 1.294, 1.099–1.523; corporate scale 0.853, 0.772–0.943; within walking distance 1.589, 1.331–1.896; continuous log-distance 0.872, 0.813–0.935; within-corporation hazard ratio 0.631, 0.477–0.835). All models are unpenalised.

Corporate identifier. Restricting every analysis to the 6,949 stations with a directly reported corporate number gave larger ownership and proximity estimates, while service scope lost significance (Supplementary Table S11); estimates without the linkage of Supplementary Methods S1 are in Supplementary Table S19.

Provider form and portfolio. Restoring the 1,254 deemed-designated medical institutions and controlling for provider form did not change the direction or significance of the organisational associations; provider form was itself the strongest single predictor in the pooled cohort (TR=0.302, 95% CI 0.269–0.340; Supplementary Table S3). Among the 7,793 stations whose corporation ran two or more

service types, ownership kept its association but service scope and corporate scale did not (Supplementary Table S4).

Urban and rural municipalities. Splitting municipalities at the median number of LTCI facilities, the continuous log-distance time ratio was 0.893 (95% CI 0.814–0.978; n=3,088) in urban and 0.911 (95% CI 0.832–0.998; n=3,176) in rural municipalities (Table 4 of the main text).

Corporations running only nursing and care management. In the most demanding subsample, the 1,143 stations whose corporation ran home-visit nursing and care management and nothing else (152 closures), the proximity association pointed the same way without reaching conventional significance (within walking distance TR=1.402, 95% CI 0.867–2.267; continuous log-distance TR=0.888, 95% CI 0.759–1.039). Service scope is constant in this subsample, so the models adjust for corporate scale and the vintage proxy only. With so few closures the subsample has little power, and the estimates are reported for completeness.

Early exits. The sensitivity of the primary estimate to the duration assigned to the 198 stations that exited before the first follow-up wave is reported in §2.3 of the main text.

Supplementary Table S14. Discrete-time survival model and closure rates by interval

Complementary log-log model on the station-wave panel, with the log of interval length as an offset, period dummies for the baseline hazard, prefecture fixed effects and standard errors clustered on the corporation. Hazard ratios below 1 indicate a lower rate of closure.

Model and covariate	Hazard ratio	95% CI	p
All stations (76,102 station-periods; 2,141 closures)			
Owns a care-management office	0.720	0.615–0.844	<0.001
Service scope (per log service type)	0.832	0.721–0.960	0.012
Corporate scale (per log office)	1.128	1.032–1.234	0.008
Agency-vintage proxy (highest vs lowest rank)	1.808	1.529–2.139	<0.001
Owners with usable coordinates (41,124 station-periods; 916 closures)			
Within 0.5 km of own care-management office (separate model)	0.723	0.623–0.840	<0.001
Log-distance to own care-management office (separate model)	1.093	1.030–1.160	0.003

The proximity models are also adjusted for service scope, corporate scale and the vintage proxy. Clustering on the 1,247 municipalities instead of the corporation gave a 95% CI for the within-0.5 km hazard ratio of 0.623–0.840, identical to the corporate-clustered interval to three decimal places, and for the log-distance hazard ratio of 1.036–1.153.

Closure rates by interval. Because only the December waves were used after 2022 (the June waves published from 2023 onwards were not), the last three intervals are twice as long as the first four.

Crude annual closure rates (primary definition), computed as closures divided by stations at risk and by interval length:

Interval	Length (years)	Stations at risk	Closures	Annual rate, %
Dec 2020 to Jun 2021	0.50	11,778	151	2.6
Jun 2021 to Dec 2021	0.50	11,580	244	4.2
Dec 2021 to Jun 2022	0.49	11,269	261	4.7
Jun 2022 to Dec 2022	0.51	10,946	240	4.3
Dec 2022 to Dec 2023	1.00	10,669	441	4.1
Dec 2023 to Dec 2024	1.00	10,144	369	3.6
Dec 2024 to Dec 2025	1.00	9,716	435	4.5

The annual rate was lowest in the first half-year (2.6%) and between 3.6% and 4.7% thereafter, with no rise in the last three intervals; the larger counts of closures at those waves reflect their longer intervals.

Supplementary Table S15. Ownership and proximity together, the binary placebo and tests of referral steering

Weibull accelerated failure time models with prefecture fixed effects, adjusted for service scope, corporate scale and the vintage proxy. A time ratio above 1 means longer expected survival.

Model	Exposure	Time ratio	95% CI	p
Three groups, n=11,757	Own office 0.5 km or more away (vs no office)	1.144	0.907–1.443	0.254
Three groups, n=11,757	Own office within 0.5 km (vs no office)	1.632	1.352–1.969	<0.001
Three groups, same model re-referenced	Own office within 0.5 km (vs 0.5 km or more)	1.426	1.194–1.702	<0.001
Placebo, owners only, n=6,264	Another operator’s office within 0.5 km	0.929	0.777–1.111	0.421
Joint model, n=6,264	Own office within 0.5 km	1.446	1.215–1.721	<0.001
Joint model, n=6,264	Another operator’s office within 0.5 km	0.943	0.788–1.127	0.517
Steering test 1, n=6,264	Proximity × centred log competitors	0.892	0.759–1.049	0.167
Steering test 2, lowest tertile of competitors (median 6), n=2,221	Own office within 0.5 km	1.559	1.161–2.094	0.003
Steering test 2, middle tertile of competitors (median 19), n=1,989	Own office within 0.5 km	1.710	1.249–2.340	<0.001
Steering test 2, highest tertile of competitors (median 41), n=2,054	Own office within 0.5 km	1.158	0.869–1.543	0.316
Steering test 3, n=6,264	Proximity × corporation’s share of local agencies	1.655	0.672–4.078	0.273

The three-group model excludes the 21 stations whose corporation ran a care-management office without usable coordinates. Being within 0.5 km of another corporation's care-management office applied to 4,379 of the 6,264 owners' stations. Referral steering predicts a larger proximity advantage where more competing agencies' clients could be captured; the interaction is below 1, the tertile estimates are smallest where competition is greatest, with overlapping intervals, and the market-share interaction is too imprecise to be informative.

Supplementary Note S16: Validation of closures, the reporting threshold and quantitative bias analysis

Reporting threshold. Providers are obliged to report to the prefecture when they start to provide services and then each year, unless they received no more than 1,000,000 yen in long-term care insurance payments in the year before the reference date, in which case reporting is voluntary (Long-Term Care Insurance Act, 1997: Article 115-35(1); Ordinance for Enforcement of the Long-Term Care Insurance Act, 1999: Article 140-44). A station whose clients are mainly paid by health insurance can therefore fall below the threshold and disappear from the open data while still operating. Of the 11,778 stations, 194 (1.6%) disappeared at one or more waves and later reappeared; these are treated as reporting gaps, not closures (§2.3 of the main text). Such gaps were more common among stations whose corporation ran no care-management office (2.3%) than among owners (1.1%), so permanent non-reporting, which cannot be distinguished from closure, may also be more common among non-owners.

Validation against published facility pages. The random sample of 800 cohort stations used to validate the vintage proxy (Supplementary Methods S2) included 146 stations counted as closures under the primary definition. When the pages were retrieved in September 2026, 144 of these (98.6%) no longer had a published page and two did, giving an observed false-closure rate of 1.4% (exact upper 95% confidence limit 4.9%). A station that continues to operate without reporting would also lack a page, so this check bounds reappearance but not silent non-reporting.

Quantitative bias analysis. To examine how false closures could affect the estimates, closures in the comparison group were reclassified at random as survivors, censored at five years, at the observed rate and at its upper confidence limit. All misclassification was placed in the comparison group (non-owners for the ownership estimate; owners whose office lay 0.5 km or more away for the proximity estimate), which is the direction that would most inflate the estimates. Each scenario was replicated 40 times.

Estimate	False-closure rate	Closures reclassified	Median time ratio	2.5th–97.5th percentile	Median lower 95% limit	Replicates with p<0.05
Care-management ownership (1.462 unadjusted for bias)	Observed (1.4%)	17	1.435	1.421–1.448	1.194	100%
Care-management ownership	Upper limit (4.9%)	59	1.373	1.354–1.392	1.140	100%
Within 0.5 km (1.449 unadjusted for bias)	Observed (1.4%)	4	1.426	1.422–1.432	1.198	100%
Within 0.5 km	Upper limit (4.9%)	16	1.357	1.344–1.365	1.138	100%

Both estimates remain significant in every replicate at either rate. The scenarios are deliberately unfavourable: They assume that every false closure falls in the comparison group.

Tipping-point analysis. Because the page check cannot detect stations that continue to operate without reporting, the same procedure was repeated at higher rates (10 replicates each) to find the share of the comparison group's closures that would have to be false for each estimate to lose significance.

Estimate	Share of comparison-group closures reclassified	Closures reclassified	Median time ratio	Median lower 95% limit	Replicates with p<0.05
Care-management ownership	10%	122	1.279	1.060	100%
Care-management ownership	15%	183	1.191	0.984	0%
Care-management ownership	20%	244	1.106	0.912	0%
Care-management ownership	25%	305	0.985	0.809	0%
Within 0.5 km	2%	6	1.417	1.189	100%
Within 0.5 km	3%	10	1.395	1.169	100%
Within 0.5 km	4%	13	1.376	1.154	100%
Within 0.5 km	6%	19	1.334	1.120	100%
Within 0.5 km	8%	26	1.298	1.089	100%
Within 0.5 km	10%	32	1.267	1.059	100%
Within 0.5 km	15%	48	1.175	0.982	0%
Within 0.5 km	20%	65	1.080	0.900	0%

Both estimates lose significance between 10% and 15%: for proximity, when between 32 and 48 of the 323 closures among owners whose office lay 0.5 km or more away are in fact stations still operating, and for ownership, when between 122 and 183 of the 1,219 closures among non-owners are. Both thresholds assume that all misclassification falls in the comparison group, and each rate was replicated only 10 times. Silent non-reporting on that scale cannot be excluded with these data, which is reported as a limitation in §4.6 of the main text.

Supplementary Table S17. Model fit across proximity thresholds, and estimates by distance band and street address

Weibull accelerated failure time models with prefecture fixed effects, adjusted for service scope, corporate scale and the vintage proxy, on the 6,264 stations whose corporation ran a care-management office with usable coordinates (916 closures). Each row replaces the binary proximity indicator with a different threshold; the continuous log-distance model is shown for comparison.

Threshold	Stations inside	Time ratio	95% CI	p	AIC	Difference from lowest AIC

Threshold	Stations inside	Time ratio	95% CI	p	AIC	Difference from lowest AIC
0.05 km	3,743	1.276	1.083–1.502	0.004	8029.09	9.85
0.1 km	4,074	1.332	1.129–1.573	<0.001	8025.94	6.70
0.2 km	4,285	1.420	1.199–1.683	<0.001	8020.35	1.11
0.5 km (primary)	4,497	1.449	1.218–1.725	<0.001	8019.24	0.00
1 km	4,683	1.462	1.215–1.759	<0.001	8019.46	0.22
2 km	4,951	1.424	1.162–1.746	<0.001	8024.06	4.82
5 km	5,370	1.288	0.996–1.664	0.054	8033.13	13.89
Continuous log(km+1)	6,264	0.903	0.843–0.967	0.004	8028.20	8.96

The 0.5 km threshold has the lowest AIC; the 1 km and 0.2 km thresholds lie within 0.22 and 1.11 units of it, too small a difference to prefer one over another, and the other thresholds and the continuous measure fit less well (differences of 4.82 to 13.89 units). Every threshold from 0.05 to 2 km gives a confidence interval excluding 1; at 5 km the interval just includes it. The threshold was chosen on the data, as §2.4 of the main text states; it is retained because it corresponds to a few minutes on foot, and the confidence intervals for the primary indicator do not account for the selection.

Distance bands and street address. Three further models on the same stations replace the binary indicator with distance bands, with owners whose office lay 0.5 km or more away as the reference (1,767 stations; five-year closure 18.3%). The second splits the band from 50 m to 0.5 km by whether the station and the office share a street address: the town, block and lot after the normalisation of Supplementary Note S21, ignoring building, room and facility names and the prefixes 大字 and 字. The third pools the stations within 50 m or at the same street address, as an approximation to sharing a site.

Model	Group	Stations	Five-year closure, %	Time ratio	95% CI	p
Distance bands	Under 50 m	3,743	13.2	1.437	1.199–1.721	<0.001
Distance bands	50 m to 0.5 km	754	13.0	1.513	1.148–1.992	0.003
Distance bands and street address	Under 50 m	3,743	13.2	1.434	1.197–1.718	<0.001
Distance bands and street address	50 m to 0.5 km, same street address	323	9.0	2.410	1.557–3.729	<0.001
Distance bands and street address	50 m to 0.5 km, different street address	431	16.0	1.147	0.832–1.582	0.402
Pooled co-location	Within 50 m or at the same street address	4,066	12.9	1.489	1.246–1.780	<0.001

Model	Group	Stations	Five-year closure, %	Time ratio	95% CI	p
Pooled co-location	50 m to 0.5 km, different street address	431	16.0	1.146	0.831–1.581	0.404

Of the 3,743 stations within 50 m of the office, 3,382 (90.4%) share its street address, and 4,066 of the 4,497 stations within 0.5 km (90.4%) are within 50 m of the office or at its street address. The walking-distance contrast of the main text therefore compares mostly stations that share a site with the office, as far as addresses can show it, with owners further away. Stations 50 m to 0.5 km away at the same street address have the largest estimate, a discrepancy that can arise from imprecise coordinates. For the stations 50 m to 0.5 km away at a different street address, the group most likely to occupy separate premises, the estimate is smaller and imprecise, and it does not differ significantly from that for stations within 50 m (ratio of time ratios 1.250, 95% CI 0.916–1.706, p=0.159) or from that for the pooled co-located group (1.299, 95% CI 0.954–1.769, p=0.097; Wald tests with the corporation-clustered covariance of the same model). Whether the advantage extends to stations on separate premises therefore cannot be determined. A street address is only a proxy for premises: Stations at the same street address may occupy separate buildings, and stations at different addresses may be neighbouring buildings of one site. Differences in AIC in the first table are computed from the rounded values shown.

Supplementary Table S18. Net change in providers by municipal characteristics, December 2020 to December 2025

Municipalities with at least one home-visit nursing provider in December 2020 (n=1,501). Providers include deemed-designated hospitals and clinics. Provision per 10,000 residents aged 75 or over uses the 2020 Census population as the denominator at both dates.

Group	Municipalities	Net loss, %	Left with no provider, %	Median change in providers, %	Providers per 10,000 aged 75+, 2020 to 2025
By share of residents aged 75 or over					
Lowest quartile	376	6.4	0.5	49	8.1 to 11.9
Second quartile	375	9.9	0.3	33	7.4 to 10.2
Third quartile	375	16.8	2.7	0	6.2 to 7.7
Highest quartile	375	28.5	8.3	0	6.0 to 5.9
By number of providers at baseline					
Fewest (median 1)	543	14.4	8.1	0	
Second (median 3)	239	23.4	0.0	0	
Third (median 7)	361	21.1	0.0	25	
Most (median 20)	358	5.9	0.0	40	

The groups defined by baseline provider count are unequal in size because many municipalities share the same count. Change in population over the period was not examined.

Supplementary Table S19. Estimates with the corporate identifier of the submitted version

In the submitted version, a record without a national corporate number was identified by prefecture and registered name even when other records of the same corporation reported the number, so that one corporation could appear as two (Supplementary Methods S1). Linking such records changed the classification of 1,209 stations (10.3% of the analytic cohort) from non-owners to owners, and of none in the opposite direction; it changed the number of service types for 2,981 stations, the number of offices for 4,037, and the walking-distance classification for 218 of the stations classified as owners under both identifiers. The classification of exits as having a provider of the same corporation nearby also uses the identifier; with the linkage, 11 stations changed closure status. The table compares estimates under the closure definition of the main text, with unpenalised models throughout.

Estimate	Identifier of the submitted version	Linked identifier (main text)
Stations whose corporation ran a care-management office	5,076	6,285
Closures (primary definition)	2,152	2,141
Owns a care-management office, TR	1.320 (1.111–1.568)	1.462 (1.217–1.756)
Service scope, TR	1.358 (1.153–1.601)	1.235 (1.048–1.456)
Corporate scale, TR	0.859 (0.771–0.958)	0.870 (0.785–0.964)
Agency-vintage proxy, TR	0.499 (0.412–0.604)	0.505 (0.415–0.613)
Owners with usable coordinates; within 0.5 km	5,059; 3,511	6,264; 4,497
Within walking distance (owners), TR	1.277 (1.054–1.546)	1.449 (1.218–1.725)
Continuous log-distance, TR	0.940 (0.870–1.016), p=0.120	0.903 (0.843–0.967), p=0.004
Own office ≥0.5 km vs none, TR	1.143 (0.905–1.443)	1.144 (0.907–1.443)
Own office <0.5 km vs none, TR	1.404 (1.173–1.680)	1.632 (1.352–1.969)
Another operator's office within 0.5 km (placebo), TR	1.036 (0.851–1.262)	0.929 (0.777–1.111)
Within-corporation comparison, HR	0.671 (0.479–0.940); 1,824 stations, 496 corporations	0.665 (0.500–0.883); 2,393 stations, 650 corporations
Correlation of log scope and log scale	0.78	0.76

Time ratios (TR) and hazard ratios (HR) with 95% confidence intervals, estimated as in the main text, with standard errors clustered on the corporation under each identifier. The linkage increases the ownership and proximity estimates, reduces the service-scope estimate and makes the continuous distance gradient significant; the estimate for owners further away relative to non-owners is unchanged, and the placebo remains close to one.

Supplementary Note S20: Exposures as time-varying covariates

Construction. The analysis uses the 8,493 stations whose corporation was identified by a national corporate number, reported or linked (Supplementary Methods S1), because only a number identifies a

corporation reliably across waves. Each station keeps its baseline corporation. At every wave, care-management offices were assigned to a corporation by the number on their record or, where none was reported, by the number reported for the same provider number at another wave or by a prefecture and normalised name that corresponded to a single number among the home-visit nursing and care-management records of all eight waves. For each period, exposure was measured at the wave that opens it: no care-management office in the corporation, its nearest office 0.5 km or more away, or within 0.5 km; a nearest office more than 1,000 km away was treated as a coordinate error and the period coded as having no office. Service scope, corporate scale and the vintage proxy are kept at their baseline values.

Change during follow-up. At the first wave the time-varying classification agreed with the baseline classification of the main analysis for 97.8% of stations. Over 55,014 station-periods, 1,087 stations (12.8%) changed category at least once, and 6.8% of station-periods were in a category different from baseline.

Baseline category	Category during the period	Station-periods
No own office	Own office ≥0.5 km	514
No own office	Own office <0.5 km	779
Own office ≥0.5 km	No own office	480
Own office ≥0.5 km	Own office <0.5 km	522
Own office <0.5 km	No own office	660
Own office <0.5 km	Own office ≥0.5 km	806

Results. Discrete-time complementary log-log models as in Supplementary Table S14, with standard errors clustered on the corporation; hazard ratios below 1 indicate a lower rate of closure.

Comparison	Exposure fixed at baseline	Time-varying exposure
Own office <0.5 km vs none	0.640 (0.528–0.777)	0.471 (0.393–0.563)
Own office ≥0.5 km vs none	0.890 (0.706–1.123)	0.712 (0.574–0.884)
Owners only: own office <0.5 km vs ≥0.5 km	0.686 (0.586–0.803)	0.630 (0.535–0.741)
Station-periods (closures), three groups	55,014 (1,476)	55,014 (1,476)
Station-periods (closures), owners only	35,208 (785)	35,361 (746)

Among owners, the proximity contrast changed little when exposures were updated. Relative to non-owners, both groups of owners fared better with time-varying exposures, and the contrast for owners further away became significant. This is what would be expected if a corporation in difficulty closes its care-management office before the station: The station then spends its last periods as a non-owner, which concentrates closures in that category. The time-varying estimates are therefore not cleaner than the fixed ones, and the main text keeps exposures fixed at baseline (§4.6 of the main text).

Supplementary Note S21: Coordinates and error in the distance measure

The source does not document how facility coordinates are assigned. The checks below examine their precision and the consequences of error for the proximity estimates.

Recorded precision. Among the 53,011 home-visit nursing and care-management records of December 2020, coordinates are recorded with between 0 and 9 decimal places, most often seven (76.2%) or nine (21.4%); 1,312 records (2.5%) share their exact coordinates with a record at a different normalised address, which suggests that some coordinates are assigned at a coarser level than the address.

Pairs registered at the same address. Of the 6,264 pairs formed by a station and its corporation's nearest care-management office, 2,972 share the same address after normalisation (Unicode normalisation, removal of spaces and of the prefecture and municipality, and a common notation for block and lot numbers; building names were not compared). If coordinates were exact, these pairs would lie at the same point. The median distance between them is zero and the 90th percentile 53 m, but 305 pairs (10.3%) are 50 m or more apart and 94 (3.2%) 0.5 km or more apart, with a maximum of 878 km, which can only be a coordinate error.

Distance between station and office	Pairs	Same address	Identical recorded coordinates
Under 50 m	3,743	2,667	1,708
50 m to under 500 m	754	211	0
500 m or more	1,767	94	0

Treating every same-address pair as co-located moves 94 stations into the walking-distance group and gives a time ratio for proximity of 1.469 (95% CI 1.227–1.760), against 1.449 in the main analysis; in the three-group model the estimates are 1.624 (95% CI 1.347–1.958) for owners within 0.5 km and 1.130 (95% CI 0.891–1.434) for owners further away, relative to non-owners.

Same and different addresses within walking distance. Of the 4,497 stations within 0.5 km of their corporation's nearest care-management office, 2,878 were registered at the same normalised address and 1,619 at a different one. Five-year closure was 13.1% and 13.4% respectively, against 18.3% among owners further away, and in one model with owners further away as the reference, adjusted as in the main analysis (n=6,264, 916 closures), the time ratio was 1.463 (95% CI 1.211–1.769) for the same address and 1.427 (95% CI 1.155–1.763) for a different address. This comparison does not separate stations that share a site from those that do not: 1,076 of the 1,619 stations at a different address were within 50 m of the office, and 715 of these shared its town, block and lot, differing only in the prefixes 大字 or 字 or in the text after the lot number, such as building, room or facility names. Supplementary Table S17 therefore separates distance from address. There, the 431 stations 50 m to 0.5 km from the office at a different street address have a smaller but imprecise estimate that does not differ significantly from that for the other stations within 0.5 km, most of which share a site with the office as far as addresses can show it; whether the advantage extends to separate premises cannot be determined.

Simulated positional error. Every distinct location of a station or care-management office was displaced by a random error drawn independently for each axis from a normal distribution with standard deviation 25, 50, 100 or 200 m; locations with identical coordinates were given the same displacement, so that offices recorded at the same point stay together. A fifth scenario left 97% of locations unchanged and displaced 3% with a standard deviation of 1 km, mimicking the share of same-address pairs found 0.5 km or more apart. Distances were then recomputed and the models refitted (20 replicates per scenario).

Scenario	Stations reclassified (median)	Within 0.5 km vs further, TR (median lower–upper 95% limit)	Replicates with p<0.05	Own office ≥0.5 km vs none, TR (median)	Replicates with p<0.05
SD 25 m	15	1.441 (1.207–1.718)	100%	1.150	0%
SD 50 m	30	1.430 (1.200–1.702)	100%	1.156	0%
SD 100 m	70	1.431 (1.200–1.707)	100%	1.156	0%
SD 200 m	628	1.289 (1.092–1.521)	95%	1.252	70%
3% displaced,	125	1.405 (1.184–1.665)	100%	1.174	0%

Scenario	Stations reclassified (median)	Within 0.5 km vs further, TR (median lower–upper 95% limit)	Replicates with p<0.05	Own office ≥0.5 km vs none, TR (median)	Replicates with p<0.05
SD 1 km					

Simulated detour. Straight-line distances were multiplied by a detour factor: a constant 1.2, 1.4, 1.6 or 2.0; factors drawn from gamma distributions with a larger mean in rural than in urban municipalities (moderate: means 1.3 and 1.6; strong: 1.4 and 2.0); or a heavy-tailed factor, one plus a lognormal variable with median 0.3. The constant scenarios are deterministic; the others were replicated 20 times.

Scenario	Stations reclassified (median)	Within 0.5 km vs further, TR (median lower–upper 95% limit)	Replicates with p<0.05
Constant 1.2	49	1.422 (1.195–1.692)	100%
Constant 1.4	81	1.397 (1.175–1.660)	100%
Constant 1.6	113	1.407 (1.185–1.669)	100%
Constant 2.0	166	1.412 (1.190–1.676)	100%
Rural > urban, moderate	87	1.400 (1.178–1.664)	100%
Rural > urban, strong	120	1.420 (1.196–1.687)	100%
Heavy-tailed	89	1.407 (1.185–1.672)	100%

Across all scenarios the median time ratio for proximity ranged from 1.289 to 1.441, and it was significant in every replicate except one of the 20 with a standard deviation of 200 m. Errors of the size suggested by the same-address check (up to about 100 m, or a few large errors) reclassify few stations and change the estimate little. With a standard deviation of 200 m, many co-located pairs are pushed beyond 0.5 km: The proximity estimate falls to 1.289, and owners further away then outlast non-owners (1.252, significant in 14 of 20 replicates) because that group now contains stations that share a site with their office. Street-network distances and travel times were not computed (§4.6 of the main text).