



Effet des radars fixes sur l'accidentalité à leurs abords : ce que disent les chiffres (2019-2024)

Étude N°02 Août 2026 France métropolitaine, période 2017-2024

Cette étude mesure l'effet local des 1 521 radars fixes datés d'une mise en service entre 2019 et 2024 en France métropolitaine, sur le nombre et la gravité des accidents corporels survenus à moins de 500 mètres puis de 1 kilomètre de la cabine. Nous concluons que les radars fixes n'ont pas d'effet significatif ni sur la mortalité, ni sur le nombre de blessés ni sur l'accidentalité en général.

Résumé

Ce que l'étude mesure. Un radar fixe fait-il baisser le nombre d'accidents et leur gravité juste à côté de lui ? Nous avons examiné les 1 521 radars fixes que le fichier officiel date d'une mise en service entre 2019 et 2024 en France métropolitaine, et les accidents corporels enregistrés à moins de 500 mètres puis de 1 kilomètre de chacun d'eux, de 2017 à 2024. L'étude porte sur cet effet **local**.

Ce que l'on voit d'abord. En comparant simplement le nombre d'accidents avant et après l'installation, la baisse est nette : environ 30 % d'accidents en moins autour des radars récents. Le chiffre est exact, mais n'est pas la conséquence de l'installation des radars fixes.

D'où vient cette baisse ? Plusieurs mécanismes, étrangers au radar, conduisent à une baisse de l'accidentalité.

Les radars ne sont pas posés au hasard : ils le sont là où les accidents se sont accumulés. Or un endroit qui vient de connaître une mauvaise série revient de lui-même vers sa moyenne habituelle, radar ou pas.

L'accidentalité baisse par ailleurs depuis des décennies sur l'ensemble du réseau, pour des raisons qui tiennent aux véhicules, aux infrastructures et aux comportements. Notre étude établit qu'une portion de route équipée en 2021 aurait connu, sans radar, l'amélioration de fond qu'ont connue les portions comparables.

Un site donné enregistre en outre peu d'accidents par an. À ces effectifs, une année calme succède souvent à une année chargée sans que rien n'ait changé sur la route.

Enfin, les radars ont été installés à des dates échelonnées entre 2019 et 2024 : les périodes « avant » et « après » n'ont donc pas la même durée d'un site à l'autre, et la façon de tenir compte de ces durées inégales pèse, à elle seule, sur le chiffre obtenu.

Ces causes se mélangent, et nos données ne permettent pas de dire laquelle pèse le plus. Restait à vérifier si notre propre méthode sait démêler ce qui revient au radar.

Comment nous avons vérifié notre méthode. Nous avons soumis nos propres calculs à un examen de contrôle : attribuer à chaque radar une **fausse date de mise en service**, deux ou trois ans avant la vraie, puis n'utiliser que les années antérieures à la vraie date de mise en service. Dans cette situation. Cet examen porte sur deux calculs. Une comparaison avant/après qui ne corrige pas la tendance déjà en cours, du même type que celle qui produit la baisse de 30 %, ne donne pas zéro : elle affiche une hausse, de 9 % à 20 % selon la façon de découper les données, nette dans six cas sur sept. C'est la démonstration directe, et simple, que l'accidentalité d'un site monte jusqu'à l'année qui précède la pose du radar. **La baisse de 30 % ne mesure donc pas l'effet du radar.** C'est le résultat le plus solide de ce travail.

Ce que l'étude conclut. Nos données ne font apparaître aucun effet détectable des radars fixes récents sur la gravité des accidents corporels ni sur la mortalité, à proximité immédiate des cabines. Ces événements sont rares, ce qui borne la sensibilité de nos données : seul un effet important, de l'ordre du quart sur les blessés graves et du tiers ou plus sur les décès, apparaîtrait dans nos chiffres. Un effet plus modeste resterait invisible, à la baisse comme à la hausse ; les seuils précis, qui diffèrent selon le sens de l'effet, sont donnés dans le résumé technique ci-dessous. Quant à la survenue des accidents, tous niveaux de gravité confondus, nous concluons que les données ne permettent pas de mesurer l'effet local des radars.

À quoi les zones équipées sont comparées, et ce qu'est une « installation ». Les zones de comparaison ne sont pas des routes sans radar : ce sont des abords de radars plus anciens, mis en service avant 2019 et non modifiés pendant la période. Et l'installation elle-même est, dans deux cas sur trois, un remplacement : le fichier des radars date l'équipement en place, non le site, et sa comparaison avec une extraction antérieure aux dégradations de l'hiver

2018-2019 montre que 1 030 des 1 521 radars datés 2019-2024 (68 %) ont pris la place d'un radar présent au même endroit en octobre 2018, le plus souvent une cabine remplacée par une tourelle. Ce que l'étude mesure est donc, pour l'essentiel, l'arrivée ou le renouvellement d'un radar dans un réseau déjà densément équipé, et non, dans deux cas sur trois, l'introduction d'un radar sur une route qui n'en avait pas. Et sur le tiers de sites qui n'avaient pas de radar auparavant ? Les données y sont plus rares (327 zones) nous ne pouvons pas davantage y mesurer l'effet local sur le nombre d'accidents ni sur la gravité et la mortalité.

Principaux résultats

Sur la gravité et la mortalité, les radars fixes installés depuis 2019 ne montrent aucun effet détectable : les ruptures estimées sont proches de zéro et non significatives (blessés graves $-1,5\%$ à 500 m et $+8,6\%$ à 1 km ; KSI, c'est-à-dire tués ou blessés graves, $-2,9\%$ et $+6,7\%$; tués -17% et $-7,5\%$), et la rareté des événements borne la sensibilité des données. Seul un effet d'au moins 18 à 25 % à la baisse (22 à 34 % à la hausse) sur la gravité, et de 33 à 48 % à la baisse (50 à 94 % à la hausse) sur la mortalité, serait décelable : la limite est ici la puissance statistique, et non l'établissement d'une absence d'effet. Sur le nombre d'accidents, nous établissons l'effet localisé n'est pas identifié de façon robuste. Les estimations disponibles sont $+3,4\%$ à 500 m ($p = 0,52$; IC 95 % $[-7\%$; $+15\%$]) et $+4,9\%$ à 1 km ($p = 0,27$; IC 95 % $[-4\%$; $+14\%$]) par test de cassure, $-3,4\%$ et $-2,0\%$ en Empirical Bayes, toutes non significatives ; elles dépendent de la manière dont on modélise le reflux d'accidentalité qui précède l'installation (§8, annexe 13.6). Les seuils de détection correspondants (environ 14 % à 500 m et 11 % à 1 km à la baisse, 16 % et 13 % à la hausse) mesurent la sensibilité de cette spécification ; ils ne borneraient l'effet réel que si celle-ci était non biaisée, hypothèse que le test placebo met en défaut (§9, annexe 13.7). En revanche, un point est établi sans dépendre d'aucune hypothèse sur la forme de la tendance : la baisse de 30 % observée en comparaison brute avant/après ne mesure pas l'effet des radars (annexe 13.7).

Statut des estimateurs et du test placebo. Le test de cassure porte l'estimation : il sépare une pente continue, estimée conjointement sur les périodes antérieure et postérieure à la pose, du saut discret lié à l'installation. L'estimation Empirical Bayes (méthode de Hauer), calibrée faute de référentiel externe adapté sur la période « avant » des sites traités eux-mêmes (poids médian $w \approx 0,05$ sur la fréquence), reste par construction exposée à la part de retour à la moyenne commune à l'ensemble du groupe traité : elle se lit comme une **borne basse** et comme un diagnostic de cohérence, non comme une seconde correction indépendante (§3.7, annexe 13.4). Le **test placebo** (annexe 13.7) classe les spécifications sans en certifier aucune : la spécification publiée s'écarte de zéro de 19,2 points en médiane, toujours dans le même sens, significativement dans trois configurations sur sept, et cet écart ne peut pas être retranché de l'estimation principale, les échantillons et les supports différant.

Méthodologie. Nous comparons les zones à radar récent (installés en 2019 ou après) à des zones témoins qui sont elles-mêmes des abords de radars, mis en service avant 2019 et pour 1 594 sur 1 646 non modifiés pendant la période (§7.2), à deux rayons (500 m et 1 km). Les

zones témoins servent à capter ce qui relève de la tendance nationale et non du nouveau radar. La correction des mécanismes décrits ci-dessus repose sur un **test de cassure en régression de Poisson**, qui sépare la tendance déjà en cours à la date de la pose de la rupture que celle-ci provoquerait ; c'est lui qui porte l'estimation. Une seconde lecture par **Empirical Bayes** (méthode de Hauer, référence internationale du domaine) sert de diagnostic de cohérence. Nous avons enfin soumis notre propre méthode à un test placebo, en attribuant aux radars une fausse date de mise en service (§8, annexes 13.6 et 13.7). La grandeur estimée est étroite, à deux titres. Les zones témoins étant elles-mêmes des abords de radars anciens, l'étude mesure l'effet marginal d'un radar de plus dans un réseau déjà densément équipé, et non l'effet du contrôle automatisé lui-même. Et pour 68 % des radars du groupe traité, l'événement daté par le fichier officiel n'est pas une première installation mais le remplacement d'un radar présent au même endroit depuis 2005-2012, souvent une cabine remplacée par une tourelle : leur période « avant » était déjà sous radar (§5, §11, annexe 13.6)

Données. ONISR (fichiers BAAC), 2017 à 2024, en prenant 2017 et 2018 comme années de référence, ainsi que le fichier des radars automatiques (état actuel, et extraction d'octobre 2018 pour la nature du traitement). Les fichiers BAAC ne recensent que les accidents **corporels**, c'est-à-dire ayant fait au moins une victime tuée ou blessée ; les accidents purement matériels n'y figurent pas et ne sont donc pas mesurés ici.

CHIFFRES CLÉS

415 623	accidents géolocalisés constituant la base (métropole, 2017-2024), dont 44 815 à moins de 500 m d'un radar, tous millésimes confondus (9 761 aux abords des seuls radars récents étudiés)
1 521	radars fixes datés 2019-2024 par le fichier des radars, dont 1 030 (68 %) remplacent un radar présent au même endroit en octobre 2018, et dont 1 002 sont le radar le plus proche d'au moins un accident survenu à moins de 500 m (groupe analysé)
Aucun effet détectable	sur les blessés hospitalisés (proxy imparfait des blessés graves), sur leur total avec les tués et sur la mortalité, à proximité immédiate des radars étudiés (limite de puissance statistique ; fenêtre 2018-2024 pour la gravité ; seuils de 18 à 25 % à la baisse sur la gravité et de 33 à 48 % à la baisse sur la mortalité, non symétriques)
Effet localisé non identifié	sur le nombre d'accidents : le signe de l'effet n'est pas déterminé par ces données (§8, annexes 13.6 et 13.7)

Le -30 % brut ne mesure pas l'effet du radar

établi sans hypothèse sur la forme de la tendance, par test placebo (annexe 13.7)

Table des matières

Résumé

Principaux résultats

1. Objectif et portée

2. Contexte historique du contrôle automatisé

- 2.1 Genèse (2002-2003)
- 2.2 Montée en puissance du parc
- 2.3 Typologie des dispositifs
- 2.4 Recettes et affectation
- 2.5 Controverses scientifiques
- 2.6 Vandalisme et « gilets jaunes » : une expérience grandeur nature
- 2.7 Perception publique
- 2.8 Portée pour le cadrage de la présente étude

3. Revue de littérature : mesurer l'efficacité des radars fixes

- 3.1 Le problème central : biais de sélection et régression vers la moyenne
- 3.2 Deux familles de protocoles à ne pas confondre
- 3.3 Hiérarchie des méthodes d'évaluation locale
- 3.4 Tableau de synthèse comparatif
- 3.5 Revue de la littérature (RTM)
- 3.6 Le cas français
- 3.7 Positionnement de la présente étude

4. Effets négatifs et externalités : revue de littérature

- 4.1 Délai de sanction et perception de rentabilité
- 4.2 Coûts économiques et sociétaux
- 4.3 Vie privée et surveillance
- 4.4 Contrôler tous pour sanctionner quelques-uns : la question de la proportionnalité

5. Données et périmètre

6. Aperçu descriptif des données

7. Choix méthodologiques

- 7.1 Le rayon d'influence : pourquoi 500 m et 1 km ?
- 7.2 Le groupe témoin : zones à radars anciens
- 7.3 La démarche en quatre étapes

8. Résultats consolidés

9. Interprétation des résultats

10. Conclusion

11. Limites

- 11.1 Ce que les données mesurent, et avec quelle précision
- 11.2 Comparabilité des groupes et identification de l'effet
- 11.3 Puissance statistique et inférence
- 11.4 Interventions concomitantes et contexte de la période

Financement, conflits d'intérêts et disponibilité des données

À propos de Réalités Routières

12. Sources et références

- 12.1 Notes de bas de page (§2, contexte historique ; §4, effets négatifs et externalités)
- 12.2 Références scientifiques et méthodologiques
- 12.3 Sources institutionnelles et données

13. Annexes

- 13.1 Sorties de régression : fréquence des accidents
- 13.2 Résultats de gravité (fenêtre homogène 2018-2024)
- 13.3 Méthodologie statistique
- 13.4 Paramètres détaillés de l'Empirical Bayes et tailles d'échantillon
- 13.5 Contrôle d'adoption échelonnée : estimateur de type Callaway-Sant'Anna
- 13.6 Sensibilité de l'estimation principale
- 13.7 Test placebo temporel

1. Objectif et portée

Mesurer si l'installation des radars fixes a une influence (à la hausse, à la baisse, ou nulle) sur **le nombre d'accidents** et sur **leur gravité** à proximité immédiate des radars. L'étude est conduite sans hypothèse de résultat : elle vise à établir ce que les données permettent de conclure, et avec quel degré de certitude. Les années **2017 et 2018** servent d'années de référence.

Portée précise. Le groupe analysé (« traité ») est constitué des radars **installés entre 2019 et 2024**, seuls à offrir une période d'observation avant *et* après leur mise en service sur la fenêtre disponible. L'étude porte donc spécifiquement sur l'effet des **radars fixes récents**. L'objet mesuré est un effet localisé, à proximité immédiate des cabines. L'effet du programme de contrôle automatisé dans son ensemble est hors du périmètre de la présente étude (§9, §10) ; il fera l'objet d'une étude dédiée.

2. Contexte historique du contrôle automatisé

Replacer la présente étude dans l'histoire du contrôle automatisé français éclaire à la fois l'objet mesuré et la façon dont son efficacité a, de longue date, été affirmée puis contestée.

2.1 Genèse (2002-2003)

Le contrôle automatisé de la vitesse procède d'un tournant politique daté. Le 14 juillet 2002, Jacques Chirac érige la sécurité routière en priorité de son second mandat, un revirement notable, alors même que les travaux préparatoires avaient en réalité été engagés dès 2001 sous le gouvernement Jospin, avec Jean-Claude Gayssot aux Transports, mais jugés trop sensibles à l'approche de l'élection présidentielle. Le premier radar est inauguré le 27 octobre 2003 sur la RN20, à La Ville-du-Bois (Essonne), et dix cabines sont en service au 1er novembre 2003.¹ Un point mérite d'emblée d'être souligné pour sa portée méthodologique : la forte baisse de la mortalité routière observée en 2002-2003 (de l'ordre de -20 %) précède largement le déploiement effectif des cabines.² Elle est donc davantage compatible avec un effet des grandes campagnes gouvernementales et médiatiques sur la sécurité routière en général, mais également un effet d'annonce et de médiatisation d'une nouvelle politique répressive, plutôt qu'un effet propre des radars : illustration précoce du fait qu'une lecture chronologique agrégée peut attribuer à un dispositif des évolutions qui lui sont, dans les faits, antérieures. Par ailleurs, traditionnellement, sous la V^e République, l'élection présidentielle était suivie d'une loi d'amnistie générale pour les infractions routières mineures. En 2002, fraîchement réélu, Jacques Chirac rompt brutalement avec cette tradition en annonçant qu'il n'y aura aucune amnistie pour les PV de la route (sauf stationnement). Son successeur, Nicolas Sarkozy, maintiendra cette rupture en 2007, aucune loi d'amnistie n'étant adoptée après l'élection présidentielle.

2.2 Montée en puissance du parc

Le parc s'est ensuite étoffé rapidement : de 257 dispositifs en 2004, il passe à 2 300 en 2008 (dont 1 473 radars fixes de contrôle de vitesse), 4 097 en 2013 et 4 428 en 2018, tous types confondus (Cour des comptes, 2021, d'après les rapports annuels de performance)³ ; on compte ensuite environ 4 530 dispositifs en 2022, puis environ 4 068 en 2025, pour un parc cible ramené à 4 160 équipements opérationnels et visibles, tous types confondus et hors radars pédagogiques (projet annuel de performances du CAS CCSR annexé au projet de loi de finances pour 2025, p. 28 à 31, voir §12.3). À cette échelle, la France demeure l'un des pays les plus densément équipés de l'Union européenne. Le cas italien, longtemps présenté

comme comptant plus de 11 000 à 13 000 dispositifs, a été fortement revu à la baisse par le recensement officiel du ministère italien des Transports (liste nationale des *autovelox* publiée le 28 novembre 2025) : sur environ 11 000 appareils informellement recensés, seuls quelque 3 150 à 3 625 se sont révélés homologués et actifs, la grande majorité ne satisfaisant pas aux exigences d'homologation.³

2.3 Typologie des dispositifs

Le terme générique de « radar » recouvre des dispositifs aux mécanismes et aux zones d'influence hétérogènes. Les **radars fixes classiques** mesurent la vitesse en un point unique depuis une cabine. Les **radars tronçon** (technologie MESTA 5000 SMART, introduite en 2012) calculent au contraire une vitesse moyenne entre deux points distants de plusieurs kilomètres. Les **radars de franchissement** contrôlent le respect des feux rouges et des passages à niveau. Les **radars mobiles ou embarqués** sont utilisés soit à l'arrêt (*mobile fixe*), soit en mouvement dans le flux (*mobile-mobile*) ; cette dernière catégorie recouvre les **voitures-radars équipées de radars mobiles de nouvelle génération (RMNG)**, dont la conduite a été externalisée à des prestataires privés (décision en 2015, expérimentation à partir de février 2017, déploiement en 2018 ; plus de 80 % de ces véhicules sont aujourd'hui opérés par le privé). Aucun de ces dispositifs n'a d'implantation fixe. À partir de 2015 apparaissent les **radars autonomes ou de chantier**, alimentés par batterie et déplaçables, puis, à partir de 2018, les **radars « tourelles »** installés sur un mât de quatre mètres, capables de contrôler simultanément la vitesse et le franchissement des feux, appelés à remplacer progressivement les cabines fixes, certaines tourelles étant d'ailleurs des leurres non équipés. S'y ajoutent enfin les **radars pédagogiques**, purement préventifs et non verbalisants, ainsi que des expérimentations récentes de radars sonores et de contrôle du covoiturage (Lyon, Strasbourg, Paris). Cette diversité justifie la prudence de la présente étude, qui traite les radars fixes de manière uniforme et n'extrapole pas ses conclusions aux autres types (voir §11, Limites). Il faut préciser d'emblée ce que cela recouvre : les installations 2019-2024 étant massivement des tourelles, l'étude porte en pratique, à 87 %, sur ce dispositif, le groupe de comparaison étant lui composé de cabines fixes, de radars de feu rouge, de radars discriminants, de tronçon et de passage à niveau (annexe 13.6, section D).

2.4 Recettes et affectation

Le contrôle automatisé génère des recettes substantielles, dont l'affectation nourrit un débat récurrent. En 2024, les amendes issues des radars ont rapporté 889 M€ (dont 658 M€ d'amendes forfaitaires et 231 M€ d'amendes majorées), en baisse de 8 % par rapport aux 965

M€ de 2023 ; sur vingt ans, le cumul avoisine 12,6 Md€. ⁴ L'engagement pris en 2003 prévoyait que l'intégralité de ces recettes serait affectée à la sécurité routière. La réalité s'en écarte : dans un rapport d'avril 2025, la Cour des comptes estime que seules 62 % des dépenses financées présentent un lien avec la sécurité routière, le reste se répartissant en 2023 entre le désendettement de l'État (environ 21 %), l'AFITF (Agence de Financement des Infrastructures de Transport en France), les collectivités, la Délégation à la sécurité routière et divers fonds, dont un fonds santé. ⁵ Les pouvoirs publics objectent que l'effort total de l'État en faveur de la sécurité routière (4,6 Md€ en 2023) excède très largement les recettes des radars, de sorte que ces dernières ne financeraient qu'une fraction de la politique publique.

2.5 Controverses scientifiques

La robustesse des évaluations officielles a été discutée dès l'origine, ce qui intéresse directement le cadrage méthodologique de la présente étude. Dès 2003-2004, le sociologue Jean-Gustave Padioleau dénonce dans *Libération*, puis dans son ouvrage *La Société du radar* (2005), l'absence d'évaluation ex-ante et d'audits indépendants ; il rapporte que ces critiques lui ont valu une suspension de ses crédits de recherche. ⁶ Les estimations de « vies sauvées » avancées au fil des ans dans le débat public se sont révélées mouvantes et contestées, sans que les évaluations censées les fonder aient pu être identifiées, avant que la communication officielle ne se fasse plus discrète sur ce point. En 2013, une étude commandée par la Ligue de défense des conducteurs (cabinet SCM) conteste le lien entre vitesse et mortalité. ⁷ À l'inverse, les évaluations de la première période du contrôle automatisé français rapportent des baisses de décès de 38 % sur autoroute, 14 % en zone urbaine et 25 à 35 % en zone rurale entre 2002 et

1. Ces valeurs portent sur le seul cas français, et non sur une estimation agrégée : le rapport ITF/OCDE de 2018, qui examine onze cas dans dix pays ayant modifié leurs limitations de vitesse ou introduit un contrôle automatisé à grande échelle, ne produit pas d'estimation d'ensemble unique. ⁸ Dans le corpus consacré aux radars, la référence favorable à l'efficacité la mieux établie est la méta-analyse de Høye (2014), qui conclut à une réduction d'environ 20 % des accidents (référence complète au §12.2). Ces positions contradictoires (des chiffres spectaculaires côté institutionnel, une contestation parfois tout aussi fragile côté critique) ont motivé la réalisation de la présente étude.

2.6 Vandalisme et « gilets jaunes » : une expérience grandeur nature

La séquence 2018-2019 offre une expérience grandeur nature rarement disponible. Le passage généralisé à 80 km/h en juillet 2018 et l'augmentation des taxes sur les carburants déclenchent une première vague de dégradations, avant que le mouvement des « gilets jaunes » n'entraîne, fin 2018, une vague de dégradations d'ampleur inédite, le ministre de l'Intérieur Christophe Castaner évaluant à environ 75 % la part des radars détruits, dégradés, attaqués ou neutralisés en mars 2019. Selon la réponse du ministre de l'Intérieur au Sénat (février 2020), le coût est estimé à environ 33 M€ de réparations en 2018 et 36 M€ en 2019, auxquels s'ajoute un manque à gagner d'environ 300 M€ lié à la destruction d'un millier de radars. Surtout, la neutralisation massive des radars s'est accompagnée d'une hausse marquée des excès de vitesse observés en décembre 2018 : signal cohérent avec un effet dissuasif du contrôle, mais à l'échelle du réseau et non à celle, localisée, que mesure la présente étude. Cet épisode a accéléré le déploiement des tourelles, réputées plus résistantes, elles-mêmes néanmoins vandalisées dès 2019 ; une nouvelle vague de dégradations accompagnera la colère agricole de 2024.⁹ Il éclaire aussi la composition du groupe étudié ici : les radars mis en service en 2019 et 2020, des tourelles pour la plupart, ont dans neuf cas sur dix pris la place d'un radar déjà présent sur le site en octobre 2018, le plus souvent une cabine installée entre 2005 et 2012 (§5, §11).

2.7 Perception publique

L'accueil du dispositif est mitigé. Les mesures engagées sous Jacques Chirac ont bénéficié d'une large adhésion initiale, malgré leur caractère coercitif : 86 % des Français approuvaient la mise en place des premiers radars automatiques en octobre 2003 et 72 % la poursuite de leur installation en octobre 2004, alors qu'en décembre 2011 l'implantation de 400 radars supplémentaires non signalés n'était plus approuvée que par 36 % d'entre eux (Ifop pour *Dimanche Ouest France*). Persiste en effet un sentiment tenace de « radar tirelire », nourri par l'impression d'une logique avant tout financière, que viennent alimenter les cas de radars « serial flasheurs » relevant jusqu'à un millier d'infractions par jour. Le même corpus d'enquêtes montre pourtant que ces deux perceptions coexistent : dans le sondage BVA de mai 2014 annexé à l'évaluation interministérielle, 71 % des conducteurs estiment que les actions publiques de sécurité routière réduisent réellement le nombre de victimes et 84 % qu'il reste possible de diminuer largement le nombre de tués.¹⁰

2.8 Portée pour le cadrage de la présente étude

L'efficacité globale du système a de tout temps été affirmée à l'aide de chiffres spectaculaires mais méthodologiquement fragiles, et contestée avec une vigueur comparable et une fragilité égale du côté de ses détracteurs. Ce déséquilibre entre l'ampleur des affirmations et la solidité des preuves justifie précisément l'approche retenue ici : une évaluation **localisée**, qui cherche à neutraliser le retour à la moyenne par un test de cassure, plutôt qu'une comparaison agrégée à l'échelle nationale, seule mobilisée par la plupart des bilans officiels ; la portée de cette correction est elle-même mise à l'épreuve au §3.7 et à l'annexe 13.7. C'est à ce type de biais (attribuer à l'installation d'un radar une baisse qui lui préexiste, comme le suggérait déjà la chute de mortalité de 2002-2003) que les corrections décrites dans les sections suivantes entendent répondre.

3. Revue de littérature : mesurer l'efficacité des radars fixes

NOTE DE POSITIONNEMENT

Cette section présente les familles de méthodes employées dans la littérature scientifique internationale pour évaluer l'efficacité des radars fixes, les résultats typiquement obtenus selon chaque approche, et le positionnement méthodologique de la présente étude au sein de ce panorama. Quatre critères ont guidé la sélection des travaux : revues systématiques et méta-analyses portant sur les radars fixes ; évaluations locales corrigeant explicitement le retour à la moyenne ; quasi-expériences publiées depuis 2024 ; variable de résultat mesurant l'accidentalité, la vitesse pratiquée étant traitée séparément (§3.4). Cette revue n'est pas systématique au sens des protocoles de type PRISMA dans le sens où nous n'agrégeons pas les effets.

3.1 Le problème central : biais de sélection et régression vers la moyenne

Le point de convergence de toute la littérature méthodologiquement sérieuse est que les radars ne sont pas implantés au hasard : ils sont posés de façon **réactive**, sur des sites sélectionnés précisément parce qu'ils ont connu un pic anormal d'accidents. Or un pic extrême, qu'il soit dû à un hasard malheureux ou à un facteur transitoire, tend naturellement à être suivi de valeurs plus basses, **indépendamment de toute intervention**. Ce phénomène, la **régression vers la moyenne (RTM, pour Regression To Mean)**, est le principal facteur de confusion des évaluations de sécurité routière (Hauer, 1997) : le « triage » des sites les plus accidentogènes conduit mécaniquement tout avant/après à surestimer l'effet réel du traitement, car une partie du recul observé n'est que le reflux de fluctuations transitoires qui ne se seraient pas reproduites de toute façon.

Un point théorique important : le biais de sélection des sites et la régression vers la moyenne sont deux biais **distincts**, qui affectent différemment l'estimation de l'effet ; la littérature issue de Hauer (1997) en tire la nécessité de les traiter séparément. Park & Lord (2010) montrent par ailleurs, sur des données de vitesse, que l'ignorance de la RTM fausse non seulement l'estimation ponctuelle mais aussi l'incertitude qui l'accompagne (erreur-type et intervalle de confiance), ce qui peut faire conclure à tort à une significativité statistique.

Les biais récurrents dans la littérature sont au nombre de quatre :

1. **Régression vers la moyenne** : reflux naturel après un pic de sélection.

2. **Maturation** : tendance de fond à la baisse (ou à la hausse) de l'accidentalité sur l'ensemble du réseau.
3. **Évolution du trafic** : un trafic en baisse réduit mécaniquement l'exposition au risque.
4. **Co-interventions** : autres mesures concomitantes (sécurité passive des véhicules, campagnes de prévention, permis à points, etc.).

La hiérarchie des méthodes présentée ci-dessous correspond largement à leur capacité à neutraliser ces quatre biais. La relation est générale, mais non universelle : **les méthodes qui neutralisent le mieux la RTM produisent des estimations plus faibles que les avant/après naïfs**, sans que l'effet disparaisse pour autant, et certaines évaluations qui testent explicitement la RTM ne détectent pas de contamination de leur résultat (Høye, 2014). Les écarts entre travaux corrigés restent importants, et le §3.5 en donne le détail.

3.2 Deux familles de protocoles à ne pas confondre

C'est la distinction la plus structurante pour situer un résultat dans la littérature, et elle explique une grande partie de l'écart apparent entre les évaluations locales rigoureuses et le discours institutionnel français. Ces deux protocoles **ne mesurent pas la même grandeur** et ne sont pas directement comparables.

Les **évaluations locales** (avant/après par site) mesurent le changement d'accidentalité dans un rayon donné autour de chaque radar. C'est le protocole le plus exposé à la RTM, car il conditionne directement sur des sites sélectionnés pour leur pic d'accidentalité passé. C'est le protocole de la présente étude.

Les **évaluations de programme** (séries temporelles interrompues) mesurent une rupture dans une courbe nationale ou régionale agrégée au moment du lancement ou du renforcement du dispositif. Ce sont elles qui produisent les grands chiffres cités dans le débat public français. Elles sont moins exposées à la RTM par sélection de site (elles ne conditionnent pas sur des points chauds), mais davantage vulnérables à la confusion par les tendances de fond et les interventions concomitantes.

Un résultat local modeste et un résultat national de rupture marqué peuvent donc tout à fait coexister sans se contredire.

3.3 Hiérarchie des méthodes d'évaluation locale

Du moins rigoureux au plus rigoureux.

3.3.1 Avant/après sans groupe témoin

C'est la méthode historiquement la plus répandue et la plus biaisée : on compare le nombre d'accidents ou de victimes sur N années avant et après l'installation, sur le seul site équipé, ce qui confond l'effet du radar avec la RTM et la tendance de fond.

La revue systématique de Pilkington & Kinra (*BMJ*, 2005) est la synthèse de référence : sur 14 études observationnelles retenues, il n'y a aucun essai randomisé contrôlé, et la majorité (8 sur 14) étaient des avant/après sans groupe témoin. Les effets rapportés sont à la fois impressionnants et très dispersés : réductions de 5 à 69 % pour les collisions, de 12 à 65 % pour les blessés et de 17 à 71 % pour les tués, ce qui traduit directement la part de biais non contrôlé. La revue Cochrane (Wilson et al., 2010) renonce de même à une méta-analyse quantitative, en raison de l'hétérogénéité et des limites méthodologiques des études incluses ; elle conclut néanmoins que la constance des réductions rapportées fait des radars « une intervention utile ». Les réductions qu'elle recense aux abords des radars vont de 8 à 49 % pour l'ensemble des collisions (la plupart entre 14 et 25 %), de 8 à 50 % pour les collisions corporelles et de 11 à 44 % pour les collisions mortelles ou graves ; sur des zones plus larges, de 9 à 35 %. C'est cette famille qui produit le **-30 %** brut que la présente étude documente à titre de comparaison pédagogique (§6).

3.3.2 Avant/après avec groupe témoin (DiD simple)

Amélioration directe : des sites non équipés sont suivis en parallèle pour capter la tendance générale. C'est la logique de la **différence-en-différences (DiD)** dans sa forme simple. Elle corrige la maturation sous une hypothèse qu'il faut toujours expliciter, dite des **tendances parallèles** : sans installation, sites traités et témoins auraient évolué de façon semblable. Cette hypothèse n'est pas directement testable (seules les tendances antérieures à l'installation peuvent être examinées). Et elle ne corrige **pas** la RTM, puisque les sites traités restent sélectionnés sur des critères d'accidentalité que les témoins ne partagent pas : le biais de sélection subsiste, même si la surestimation est réduite.

3.3.3 Empirical Bayes (Hauer, 1997)

Aujourd'hui reconnue comme la méthode de référence de l'évaluation avant/après en sécurité routière (Hauer, 1997 ; Hauer et al., 2002 ; Persaud & Lyon, 2007). On estime, via une *Safety Performance Function* (SPF), typiquement une régression binomiale négative calibrée sur des sites comparables non traités, la fréquence d'accidents *attendue* sur chaque site ; on la combine avec le comptage observé avant traitement pour estimer le contrefactuel « ce qui se serait passé sans radar ». La différence avec le comptage post-installation est l'effet attribué à

la mesure. Son atout est de corriger simultanément la RTM, l'évolution du trafic et les autres variations non imputables au radar ; ses limites sont d'être lourde en données et en expertise, et d'exposer au risque d'un « tampon de validité » accordé trop vite à toute étude qui l'emploie (Persaud & Lyon, 2007).

Des développements plus récents (score de propension, estimateurs doublement robustes, Full Bayes) ont été proposés en réponse aux limites de l'Empirical Bayes. Une réserve mérite d'être signalée : des travaux récents (Gallagher & Fisher, 2024) contestent que l'Empirical Bayes suffise à corriger *entièrement* le biais de RTM, en le démontrant analytiquement et par une expérience Monte-Carlo de type placebo. C'est un argument pour ne pas faire reposer une conclusion sur l'EB seul. Dans la présente étude, l'EB n'est pas employé comme seconde correction de la RTM (§3.7) : sa fonction est celle d'une borne basse et d'un diagnostic de cohérence.

3.3.4 Modèles économétriques de panel et semi-paramétriques

Plutôt que de raisonner site par site, cette famille modélise l'occurrence des accidents sur l'ensemble du territoire dans le temps, en exploitant la variabilité géographique et temporelle des installations. L'étude française la plus aboutie est celle de **Roux & Zamora** (*Économie et Statistique*, 2013), qui mesure l'effet local des radars installés à partir de 2003 sur un comptage communal (1998-2007), via un modèle semi-paramétrique estimé par la méthode des moments généralisés non linéaires (GMM) : l'effet est **important juste après l'installation, puis réduit au-delà de six mois sans disparaître** (–11 % sur les accidents, –25 % à –50 % sur les décès), chiffres établis sur les communes de moins de 6 000 habitants, les seules pour lesquelles leurs estimations sont significatives. Les auteurs interprètent ce déclin comme une accoutumance comportementale ; une composante transitoire de retour à la moyenne peut aussi y contribuer, mais cette lecture nous est propre et n'est pas la leur. À l'opposé méthodologique, deux travaux relèvent du protocole « de programme ». **Carnis & Blais** (2013) estiment environ 15 000 accidents mortels et 62 000 accidents non mortels évités entre novembre 2003 et décembre 2010. **Blais & Carnis** (2015) produisent une estimation distincte : une baisse d'environ 27 % des tués pour 100 000 habitants sur 2002-2010, soit 20 040 vies, attribuée conjointement à l'annonce publique de 2002 et à l'introduction du dispositif. Ni l'une ni l'autre ne comporte de correction de la régression vers la moyenne au niveau des sites, d'où des estimations sensiblement plus élevées que celles des méthodes localisées corrigées ; ces deux estimations portent sur l'effet global et national, non sur l'effet localisé.

3.3.5 Séries temporelles interrompues (ITS) et kink-test

On modélise la trajectoire temporelle d'un indicateur et on teste une rupture (de niveau et/ou de pente) à la date d'intervention (*interrupted time series*, ITS). L'estimateur employé ici en relève : il sépare explicitement une pente continue, estimée conjointement sur les périodes antérieure et postérieure à la pose, d'une rupture discrète à la date d'installation. Ce modèle est désigné dans la littérature par « kink-test Poisson » ; le présent rapport l'appelle **test de cassure**, et les deux termes désignent la même méthode. L'effet du radar est identifié par la rupture et non par la pente ; l'identification repose donc sur l'hypothèse que la pente estimée capte bien le reflux lié à la sélection des sites, et non une autre tendance propre aux sites traités, hypothèse que le test placebo de l'annexe 13.7 met en défaut dans la présente application. Limites principales : sensibilité au choix de la fenêtre temporelle, nécessité de contrôler les co-interventions, et sensibilité à la forme donnée à cette pente, vérifiable par test placebo.

3.3.6 Méta-analyses et revues systématiques

Les méta-analyses agrègent les estimations de multiples études primaires. Leur enjeu propre est le traitement du **biais de publication** : la tendance à ne pas publier les résultats nuls. Høye (2014) y répond, sur le corpus radars, par la méthode du *trim-and-fill* (Duval & Tweedie, 2000), qui estime et corrige les effets « manquants » du côté défavorable : sa méta-analyse conclut à une réduction d'environ 20 % du nombre total d'accidents, décroissante avec la distance au radar, et ne détecte pas de contamination de ce résultat par la RTM ou le biais de publication (sauf sur les accidents mortels). Une revue systématique très récente (Amancio et al., *Accident Analysis & Prevention*, 2026) systématise le paysage méthodologique des évaluations **portant sur la vitesse pratiquée** (procédures d'évaluation, variables de résultat, stratégies de mesure), couvrant les études publiées de 1982 à 2024 ; sa méta-analyse conclut à une baisse moyenne de 7,57 km/h aux abords des radars. Cette position méthodologique, à savoir que les méta-analyses constituent la meilleure voie d'analyse, ne vaut toutefois que si les études agrégées contrôlent elles-mêmes la RTM ; à défaut, la méta-analyse gagne en précision sans gagner en justesse.

3.4 Tableau de synthèse comparatif

MÉTHODE	RÉFÉRENCES CLÉS	VARIABLE MESURÉE	CONTRÔLE DE LA RTM	EFFET TYPIQUE RAPPORTÉ	LIMITE PRINCIPALE
Avant/après sans témoin	Corpus historique ; 8 études/14 chez Pilkington & Kinra (2005)	Accidents / blessés / tués sur site équipé, avant vs après	Nul	Très dispersé : -5 à -69 % collisions ; -17 à -71 % tués	Confond RTM et tendance → surestimation systématique
Avant/après avec témoin (DiD simple)	Forme DiD standard	Idem + sites témoins non équipés	Partiel	Réduit par rapport au sans-témoin	Neutralise la maturation (sous tendances parallèles) mais pas la sélection des sites → RTM résiduelle
Empirical Bayes (Hauer)	Hauer (1997) ; Hauer et al. (2002) ; Persaud & Lyon (2007) ; Gallagher & Fisher (2024)	Comptages observés + SPF (binomiale négative) → contrefactuel	Fort (RTM + trafic)	Modeste et plus incertain ; bénéfice retenu sur blessés, plus faible sur tués	Lourde en données ; correction possiblement incomplète de la RTM
Panel / semi-paramétrique	Roux & Zamora (2013) ; Carnis & Blais (2013) ; Blais & Carnis (2015)	Occurrence accidents par commune, base mensuelle/trimestrielle	Partiel à fort selon la spécification	Variable : effet fort puis décroissant (Roux & Zamora) vs effet global élevé (Carnis & Blais, 2013 ; Blais & Carnis, 2015)	Résultats dépendants des hypothèses de modélisation
ITS / kink-test Poisson	ITS standard ; kink-test (présente étude)	Trajectoire temporelle, rupture à la date d'installation	Fort si bien spécifié (sépare pente RTM du kink)	Dépend du site et de la fenêtre	Sensible au choix de la fenêtre et aux co-interventions
Power model	Nilsson (2004) ; Elvik (2005) ; Elvik et al. (2019)	Relation vitesse/accidents appliquée à une variation de vitesse	Contourne la question (mesure la vitesse, pas les accidents)	Effet sur les accidents déduit de la variation de vitesse, non mesuré sur l'accidentalité	Indirect : Suppose une relation vitesse/accidents, ici estimée sur d'autres corpus, principalement des changements de limitation de vitesse, et sa transposition à l'abord d'un radar est une hypothèse ; l'ampleur obtenue dépend de l'exposant retenu
Revue systématique sur la vitesse pratiquée	Amancio et al. (2026)	Vitesse pratiquée aux abords des radars, sans conversion en accidents	Sans objet : la variable de résultat n'est pas l'accidentalité	-7,6 km/h de vitesse moyenne	Ne renseigne pas directement l'accidentalité
Méta-analyse / revue systématique	Pilkington & Kinra (2005) ; Wilson et al. (2010) ; Høye (2014, trim-and-fill)	Agrégation d'estimations d'études primaires	Hérite des biais des études incluses ; traite le biais de publication	≈ -20 % d'accidents (Høye, 2014) ; niveau de preuve jugé faible chez Pilkington & Kinra	Ne vaut que ce que valent les études incluses

3.5 Revue de la littérature (RTM)

C'est le résultat le plus directement comparable à celui de la présente étude. Une fois la RTM correctement traitée, l'effet ne disparaît généralement pas, mais il se réduit nettement et devient souvent fragile sur les indicateurs de gravité (tués, blessés graves).

- **Londres** (Gorell & Sexton, 2004, TRL PPR027) : 77 sites de radars répondant au critère d'installation d'au moins quatre collisions mortelles ou graves par kilomètre au cours des trois années précédant la pose, chaque site étant défini par les 400 m de part et d'autre de la caméra, comparés à des sites témoins appariés sur l'historique de collisions, le trafic, la longueur et la localisation, ce qui sépare l'effet du radar des effets de tendance et de retour à la moyenne, avec une méthode Empirical Bayes en complément : réduction estimée de 12,4 % (IC 95 % : 4,2 à 20,6) de l'ensemble des collisions corporelles et de 20,6 % (IC 95 % : 3,8 à 37,4) des collisions mortelles ou graves, résumées à 12 % et 21 % dans les conclusions du rapport.
- **Programme national britannique** (Gains et al., 2005) : sur l'ensemble du programme (38 partenariats, plus de 4 100 sites), la baisse observée avant toute correction du retour à la moyenne atteint 22 % pour les collisions corporelles et 42 % pour les tués ou blessés graves ; sur le sous-échantillon de 216 sites urbains documenté au chapitre 4 et à l'annexe H, l'effet revient à 19 % après correction de tendance et correction Empirical Bayes, le retour à la moyenne expliquant environ un quart de la baisse observée des collisions corporelles et trois cinquièmes de celle des collisions graves.
- **Norvège** (Høye, 2015), l'étude Empirical Bayes probablement la plus proche méthodologiquement du présent projet : sur les sections de 100 m en amont à 1 km en aval des radars, réduction statistiquement significative de 22 % des accidents corporels, mais aucun résultat significatif sur les tués et blessés graves (KSI), ni sur les sections plus longues.
- **Angleterre** (Graham et al., 2019), via un estimateur bayésien approché doublement robuste : réduction moyenne estimée de l'ordre de 15 % des collisions aux sites de radars.
- **Agrégation internationale** (Høye, 2014) : méta-analyse de 63 estimations issues de 15 études sur radars fixes, concluant à une réduction d'environ 20 % du nombre total d'accidents, décroissante avec la distance au radar, et ne montrant pas de contamination par la RTM ni par le biais de publication (sauf sur les accidents mortels). C'est l'estimation agrégée la plus directement comparable au paramètre estimé ici.
- **Dayton, Ohio** (Bedsworth, Weber & Willardsen, 2025) : cette évaluation occupe une place à part, car elle n'ajuste pas le biais de sélection, elle s'en affranchit. Des décisions de justice et des conflits politiques ont successivement éteint, rallumé puis rééteint le

programme de radars autoroutiers, produisant une variation d'exposition **exogène** à l'accidentalité des sites. Après un grand nombre de tests sur différents échantillons, les auteurs ne détectent pas d'effet significatif des radars autoroutiers sur l'accidentalité. C'est, dans ce corpus, l'identification la moins dépendante d'hypothèses de modélisation ; elle ne détecte pas d'effet, sans pouvoir exclure un effet modéré.

- **New York** (Stagoff-Belfort, Ben-Menachem & Beck, 2025) : différences de différences échelonnées sur 1 868 radars de zones scolaires ; les auteurs rapportent une baisse de 5 % des collisions et de 2,5 % des personnes blessées par mois sur les sept premiers mois, mesurée dans un rayon d'environ 137 m autour du radar. Une seconde évaluation du même programme new-yorkais de zones scolaires (Gao et al., 2025), de protocole longitudinal sur plus de 1 800 radars, rapporte une baisse d'environ 14 % des accidents et de 75 % des infractions de vitesse à la fin de 2021, l'effet s'établissant en six mois ; portant sur le même parc que l'étude précédente, elle n'en constitue pas une réplication indépendante.
- **Pourquoi cette dispersion ?** Job (2022) consacre un article entier à cette question et recense les causes qui font varier les résultats d'une évaluation de radars à l'autre : correction ou non de la régression vers la moyenne, choix du groupe de comparaison, fenêtre spatiale et temporelle, définition de l'indicateur. Notons que son propos est explicitement favorable au contrôle automatisé.

En résumé, la littérature établit un effet réel mais modéré sur les accidents corporels, signal plus fragile et plus dispersé sur la mortalité et la gravité, où les intervalles sont larges, les résultats hétérogènes d'une étude à l'autre et l'effet décroissant quand la fenêtre spatiale s'élargit. Or, notre étude ne trouve pas ces résultats. Les seules baisses de mortalité que l'on pourrait trouver, celles d'un estimateur de diagnostic aux deux rayons, n'y sont corroborées ni par les blessés graves ni par le test de cassure (§8). La plupart des études qui corrigent la RTM conservent un bénéfice modéré sur les accidents corporels, de l'ordre de 12 à 22 % dans les exemples ci-dessus. La présente étude ne retrouve pas ce bénéfice sur les radars français récents, sans pouvoir établir qu'il est absent : sur la fréquence, son estimation dépend de la forme donnée au reflux antérieur à la pose, et aucune des formes testées ne restitue zéro là où l'effet vaut zéro (§3.7, annexe 13.7). Le rapport à cette fourchette n'est donc pas mesurable ici.

Il faut noter que cette fourchette n'est pas unanime, et que la dispersion s'observe aussi entre les travaux les plus récents et les mieux identifiés : deux quasi-expériences publiées en 2025, à Dayton et à New York, aboutissent l'une à l'absence d'effet significatif, l'autre à une baisse de l'ordre de 5 % des collisions par mois sur les sept premiers mois. Cette dispersion entre

travaux récents et bien identifiés interdit de traiter la fourchette de 12 à 22 % comme un consensus établi : c'est à cette dispersion, et non à une valeur centrale, que les résultats de la présente étude doivent être rapportés (§3.7).

3.6 Le cas français

Les évaluations françaises de référence relèvent soit du protocole « de programme » (rupture dans les séries agrégées nationales), soit d'un avant/après local sans correction explicite de la RTM au niveau des sites ; elles ne sont donc pas directement comparables au protocole local corrigé employé ici.

- **Méthodologie officielle d'incidence (DSCR/DSR), appliquée notamment par le CEREMA.** Le protocole officiel relève d'une **comparaison avant/après classique** : cinq ans avant et cinq ans après l'installation, en neutralisant l'année de pose, sur deux zones centrées sur le radar, 1 000 mètres et 4 000 mètres. La première traduit la proximité immédiate du radar ; la seconde englobe plus largement sa zone d'influence présumée et sert également de référence pour apprécier le niveau de mortalité observé dans la zone de 1 000 mètres : la réponse de la DSCR au rapporteur spécial du Sénat (2017) présente l'effet du radar comme ayant « permis, en termes de densité de tués, de ramener la zone de proximité immédiate (1 000 mètres) au niveau moyen d'une zone plus large (4 000 mètres) ». Les éléments publics ne permettent pas de déterminer si cette comparaison constitue un dispositif de contrôle local au sens statistique ou seulement une référence descriptive. Le CEREMA l'applique pour le compte de l'État mais le désigne lui-même comme la « méthode d'évaluation DSCR sur 5 ans » : il en est l'opérateur, non l'auteur, la méthode relevant de la Délégation à la sécurité et à la circulation routières, aujourd'hui Délégation à la sécurité routière (étude *Performance des radars fixes dans les Yvelines*, CEREMA, V01, décembre 2014, annexe 5.3 ; la phase publiée de cette étude classe les radars sur deux ans avant et après la mise en service). La circulaire du 3 février 2004, citée par le Sénat (2017), demandait déjà d'examiner « à la fois l'accidentologie très localisée, et l'accidentologie sur une section plus longue », sur une section localisée de 850 mètres et une section plus longue de 4 000 mètres : la logique d'une analyse locale doublée d'une analyse élargie précède donc le CEREMA, le rayon de proximité ayant évolué de 850 à 1 000 mètres. Cette méthodologie ne comporte pas, dans les éléments publics consultés, de correction explicite du retour à la moyenne au niveau des sites ; la comparaison spatiale décrite ci-dessus ne doit pas être assimilée sans précaution à un groupe témoin statistique. L'analogie avec le choix opéré ici (500 m retenu ; 2 000 m testé

et rejeté pour dilution du signal) n'est donc que partielle : les deux protocoles comparent une zone proche à une zone plus large, mais cette comparaison spatiale n'est pas, en elle-même, une correction du retour à la moyenne.

- **Carnis & Blais (2013, 2015).** Par séries temporelles interrompues sur le programme national : introduction du dispositif associée à des baisses respectives de 7,3 % et 20,7 % des taux mensuels de blessés non mortels et mortels pour 100 000 véhicules, avec des effets plus marqués à court terme (rendements marginaux décroissants). Carnis & Blais (2013) estiment environ 15 000 accidents mortels et 62 000 accidents non mortels évités entre novembre 2003 et décembre 2010. Blais & Carnis (2015) estiment séparément que **l'annonce publique de 2002 et l'introduction du dispositif, prises ensemble**, sont associées à une baisse d'environ 27 % des tués pour 100 000 habitants sur 2002-2010, soit 20 040 vies : les auteurs ne dissocient pas l'effet de l'annonce de celui des cabines, précision cohérente avec le constat chronologique du §2.1. Ces résultats portent sur l'effet **global et national** et ne s'appliquent pas directement à l'efficacité localisée site par site. Menées au niveau national et sans groupe témoin non traité, ces séries temporelles interrompues sont en outre exposées au **biais de variable omise** : elles ne peuvent isoler l'effet propre des radars de celui des réformes concomitantes (mobilisation nationale de 2002-2003, sécurité passive des véhicules, tendance de fond). À l'échelle agrégée, ce biais de co-intervention supplante la RTM comme principale menace d'identification.
- **Roux & Zamora (2013), *Économie et Statistique* n° 460-461, p. 37-68 :** l'évaluation localisée la plus rigoureuse pour la France à ce jour, et celle dont le protocole est le plus proche du nôtre. **Nous ne retrouvons pas leur résultat, sans pouvoir établir qu'il est absent de nos données.** Sur la fréquence, notre estimation dépend de la forme donnée au reflux antérieur à la pose (§3.7), et l'une de nos variantes, la pente estimée sur les seules années antérieures, en retrouve l'ordre de grandeur (-16,1 % à 500 m, annexe 13.6, section A) ; cette variante n'a pas pour autant le statut d'une estimation rivale de même valeur, étant sous-identifiée sur l'échantillon principal et ne restituant pas davantage zéro au test placebo. Le modèle de Roux & Zamora documente un effet décroissant dans le temps, mais qui ne s'annule pas : de -13 % à -17 % sur les accidents corporels et de -50 % à -75 % sur les décès dans les premiers mois suivant l'installation, ramené **au-delà de six mois à -11 % sur les accidents et -25 % à -50 % sur les décès**, soit de l'ordre de 740 décès, 2 750 blessés graves et 2 230 accidents évités dans les communes de moins de 6 000 habitants, principalement en rase campagne (données 1998-2007, extrapolées jusqu'en 2011). Trois différences peuvent expliquer ce désaccord de conclusion : la génération de radars considérée (première vague 2003-2007 contre installations 2019-2024), la nature de l'événement daté, ces auteurs mesurant des premières installations sur des sites jusque-là non équipés quand 1 030 des 1 521 radars de notre groupe traité

remplacent un radar présent au même endroit en octobre 2018 (§5, §11.2), et la spécification de la correction du biais de sélection. Nous ne pouvons pas trancher entre ces explications, cf. notre discussion au §3.7 et la puce « Comparabilité des groupes » du §11.

- **ITF/OCDE (2018).** Examine onze cas dans dix pays ayant modifié leurs limitations de vitesse ou introduit un contrôle automatisé à grande échelle, en incluant les travaux de Carnis & Blais et de l'ONISR ; la revue ne produit pas d'estimation agrégée unique. Les baisses de décès de 38 % sur autoroutes, 14 % en zone urbaine et 25 à 35 % en zone rurale entre 2002 et 2005 relèvent du cas français et de la phase initiale d'introduction du dispositif ; elles ne sont pas directement transposables à la période 2017-2024.
- **Reprises institutionnelles.** Une première évaluation officielle attribuée au contrôle automatisé environ 75 % de la baisse de la mortalité routière sur la période 2002-2005 ; cette estimation repose sur une application sommaire de la formule de Nilsson, c'est-à-dire sur la conversion d'une variation de vitesse moyenne en vies épargnées, et non sur une mesure directe de l'accidentalité (chiffre rapporté et discuté par Carnis & Blais, 2013). Dès le rapport annuel de performances 2008 du compte d'affectation spéciale, le responsable du programme 752 affirme que « dans les zones accidentogènes couvertes par les radars, les accidents ont quasiment disparu » (p. 42), sans présenter d'estimation corrigée du retour à la moyenne ni de groupe de comparaison. Une attribution du même ordre, mais sur une autre période, est reconduite d'année en année dans les documents budgétaires : le projet annuel de performances du compte d'affectation spéciale annexé au projet de loi de finances pour 2025 écrit que le déploiement des radars « a contribué à près des trois quarts de la baisse de la mortalité constatée entre novembre 2003 et décembre 2010 », et les réponses ministérielles aux questions écrites de 2022-2023 renvoient à « une évaluation réalisée en 2013 » portant sur la même période. Aucun de ces documents ne nomme l'étude sous-jacente, que nous n'avons pas pu identifier : le chiffre circule dans les textes officiels sans référence vérifiable. La mission d'évaluation interministérielle de 2014 conclut pour sa part à une « efficacité avérée » du contrôle sanction automatisé des vitesses, mais sans avancer d'attribution chiffrée et en notant que « la relation de cause à effet de chaque mesure n'est pas parfaitement établie ». Le rapport de la Cour des comptes de 2021 crédite le contrôle automatisé de « résultats incontestables, à un coût maîtrisé » (p. 16 et 41), sans reprendre cette décomposition.

3.7 Positionnement de la présente étude

La littérature méthodologique fait ressortir une régularité robuste : les méthodes qui neutralisent le mieux la RTM et le biais de sélection des sites produisent des estimations plus faibles que les avant/après naïfs. Cette réduction ne signifie pas disparition : la plupart des études corrigées conservent un bénéfice modéré sur les accidents corporels ; elle indique seulement la part du recul brut imputable à un artefact statistique plutôt qu'au radar.

La présente étude applique au cas français, sur la période récente (2017-2024), un **test de rupture (kink-test)** en régression de Poisson séparant une pente continue du saut discret lié à l'installation ; c'est lui qui porte l'estimation. Une lecture **Empirical Bayes** (méthode de Hauer, modèle de référence binomial négatif) l'accompagne à titre de borne basse et de diagnostic de cohérence, non de seconde correction. Ce protocole se situe dans le haut de la hiérarchie des évaluations locales, sous la réserve de spécification établie à l'annexe 13.7 ; la plupart des évaluations officielles françaises ne disposent pas de cette correction au niveau des sites.

Un point notable : une logique voisine, qui consiste à identifier l'effet hors de la trajectoire pré-existante plutôt que par comparaison au pic, a été discutée par D. Spiegelhalter (2013) à propos de l'approche de la RAC Foundation, qui écarte les trois années précédant l'installation. Ici, c'est la séparation entre la pente de reflux et le saut à l'installation qui joue ce rôle : l'effet du radar est mesuré comme un écart au prolongement de cette pente, et non comme un écart au niveau du pic. Cette pente est estimée **conjointement sur les périodes antérieure et postérieure** à l'installation ; l'identification repose donc sur l'hypothèse que le reflux se poursuit après la pose au même rythme qu'avant, hypothèse dont le §11 et l'annexe 13.6 mesurent le coût.

Les deux estimations produites ne sont pas deux corrections de la RTM équivalentes et indépendantes. Le test de cassure identifie l'effet hors de la trajectoire pré-existante. L'estimation Empirical Bayes, calibrée faute de référentiel externe adapté sur la période « avant » des sites traités eux-mêmes (poids médian $w \approx 0,05$ sur la fréquence, annexe 13.4), reste par construction exposée à la part de retour à la moyenne commune à l'ensemble du groupe traité (Gallagher & Fisher, 2024) : elle doit être lue comme une **borne basse** de l'effet et comme un diagnostic de cohérence, non comme une seconde correction indépendante. Le test placebo qui lui a été appliqué le confirme : dans une fenêtre antérieure à la pose réelle, elle trouve l'accidentalité de l'année précédant la pose supérieure de 7 à 13 % à ce que l'historique du site laisse attendre, dans les quatre configurations testées (annexe 13.6, section E). L'écart d'environ 7 points entre le -3,4 % de l'Empirical Bayes et le +3,4 % du test

de cassure sur la fréquence tient au terme de pente que seul le second inclut ; ce terme est compatible avec une RTM résiduelle comme avec toute autre tendance propre aux sites traités (renouvellement de chaussée, évolution du trafic local, effet d'anticipation), et rien dans les données ne permet de les départager (annexe 13.4).

Sur les **indicateurs de gravité et de mortalité**, le résultat obtenu (aucun effet détectable) est à comparer à une littérature corrigée dont les résultats divergent sur ce registre : Høye (2015) ne trouve aucun résultat significatif sur les tués et blessés graves, tandis que Gorell & Sexton (2004) estiment une réduction de 20,6 % des collisions mortelles ou graves et Gains et al. (2005) une réduction de 19 % des collisions graves ou mortelles après correction de tendance et du retour à la moyenne ; la méta-analyse de Høye (2014) donne une réduction de 51 % sur les accidents mortels, seul résultat dont l'auteure signale qu'il peut être affecté par le retour à la moyenne (§3.5 pour les deux premières valeurs, §12.2 pour la troisième).

Sur la **fréquence des accidents**, la littérature corrigée conserve un bénéfice modeste, de l'ordre de 12 à 22 % (§3.5), et l'évaluation française la plus proche, Roux & Zamora (2013), conclut à -11 % au-delà de six mois. Il serait commode d'écrire que la présente étude, avec +3,4 %, se situe sous cette fourchette et la contredit. Ce serait aller au-delà de ce que les données autorisent : cette estimation ponctuelle dépend de la forme donnée au reflux d'accidentalité antérieur à la pose, et le test placebo de l'annexe 13.7 montre que la spécification retenue ne restitue pas zéro là où l'effet vaut zéro, sans qu'aucune des autres formes testées se comporte assez bien pour être certifiée : les seuils de détection du placebo lui-même, de 17,7 % à 48,3 % à la baisse selon la spécification, sont du même ordre que les écarts qu'il mesure. **L'étude ne permet donc pas de situer son résultat par rapport à la littérature sur cet indicateur** ; elle permet seulement d'affirmer qu'elle ne retrouve pas, sur les radars français récents, le bénéfice localisé qu'une partie de la littérature rapporte, sans pouvoir établir que ce bénéfice est absent.

Cette prudence n'est pas isolée dans la littérature récente. L'évaluation dont l'identification est la moins dépendante d'hypothèses de modélisation, l'expérience naturelle de Dayton, où le programme a été éteint puis rallumé pour des motifs étrangers à l'accidentalité (Bedsworth et al., 2025), ne trouve pas non plus d'effet significatif, sur les radars autoroutiers d'une seule agglomération, tandis qu'une évaluation quasi-expérimentale new-yorkaise de la même année en trouve un, à court terme et à très courte distance (Stagoff-Belfort et al., 2025). La dispersion des résultats reste, à ce jour, la caractéristique première de ce champ.

Ce qui est en revanche établi ici, et sans dépendre d'aucune hypothèse sur la forme de la tendance, c'est que **l'essentiel du -30 % brut n'est pas attribuable au radar** (annexe 13.7). Il faut être précis sur ce qu'il devient : la décomposition de l'annexe 13.4 montre que cet écart provient pour environ deux tiers du passage d'une comparaison agrégée à une pondération site à site, et pour un tiers de la tendance du groupe témoin. La part strictement imputable au retour à la moyenne est identifiée par le seul terme de pente du test de cassure, et rien dans les données ne permet de distinguer cette pente d'une autre tendance propre aux sites traités.

4. Effets négatifs et externalités : revue de littérature

NOTE DE POSITIONNEMENT

*Une littérature critique et plusieurs rapports officiels documentent des **effets négatifs ou des externalités** du contrôle automatisé : psychologiques, économiques, de vie privée, et liés au délai de sanction ; s'y ajoute une interrogation de philosophie du droit sur la légitimité d'un contrôle exercé sur tous (§4.4). Ces dimensions sont présentées ici pour éclairer les limites du dispositif.*

4.1 Délai de sanction et perception de rentabilité

La **célérité** (rapidité) de la sanction est un pilier classique de la théorie de la dissuasion (Beccaria, Bentham) : plus le délai entre l'infraction et la sanction est long, plus le lien psychologique entre l'acte et ses conséquences s'affaiblit. Plusieurs mécanismes sont avancés dans la littérature. D'abord une **réduction de l'impact émotionnel et de l'apprentissage** : le cerveau associe moins fortement l'excès de vitesse à la punition lorsque celle-ci survient plusieurs jours ou semaines après. Ensuite un **escompte temporel** (*temporal discounting*) : les individus valorisent moins une punition différée qu'une sanction immédiate, de sorte qu'un procès-verbal reçu 10 à 30 jours plus tard est psychologiquement moins dissuasif qu'un contrôle par un agent. S'y ajoutent une **frustration et un sentiment d'injustice** (le conducteur ne se souvient souvent plus précisément de l'infraction au moment de recevoir l'avis) et une atténuation de la **dissuasion spécifique** chez certains conducteurs, même lorsque l'effet global reste positif.

Cette dimension temporelle se combine avec la **perception de rentabilité** du dispositif. Lorsque les conducteurs estiment que les radars servent davantage à générer des recettes qu'à améliorer la sécurité, le procès-verbal est vécu comme une « taxe déguisée » plutôt que comme une sanction légitime, ce qui peut déclencher une **réactance psychologique** (résistance à un contrôle perçu comme illégitime). Cette perception est documentée : en France, le sentiment de « radar tirelire » (§2.7) est ancien, le rapport du Sénat (2017) relève la qualification de « pompes à fric », et la Cour des comptes a critiqué la lisibilité de l'affectation des recettes (§2.4). À l'étranger, la même perception est documentée par l'évaluation quadriennale britannique, qui mesure, dans les enquêtes locales de 2002-2004 conduites sur 35 zones de partenariat, 55 % d'accord en moyenne avec l'énoncé « *cameras are an easy way of making money out of motorists* », en progression sur les éditions antérieures (45 % puis 52 %) et avec une dispersion territoriale considérable, de 15 % à 74 %. La même enquête relève cependant 82 % d'accord avec l'idée que les radars doivent être soutenus comme moyen de

réduire le nombre de victimes, 71 % avec l'idée que leur but premier est de sauver des vies, et 21 % seulement de conducteurs estimant qu'il y a trop de radars dans leur zone (Gains et al., 2005, chapitre 5). Le passage de cette perception à un affaiblissement de la légitimité perçue, puis de l'effet dissuasif de long terme, relève d'une hypothèse issue de la théorie de la justice procédurale (Tyler, 1990), et non d'un résultat mesuré. Les deux perceptions coexistent : le rapport britannique conclut à un soutien large, assorti d'une préoccupation persistante quant à la finalité financière du dispositif.

À ce titre, une sanction délivrée lors d'une interception immédiate par un agent tend à être perçue comme plus légitime qu'un avis automatisé reçu à distance : l'échange interpersonnel ménage au conducteur une possibilité d'explication, ressort central de la justice procédurale, et rattache le contrôle à un objectif de sécurité plutôt qu'à un motif de recettes (Tyler, 1990).

Ces mécanismes théoriques doivent toutefois être confrontés à l'apport empirique le plus direct disponible sur ce point. En France, les délais d'envoi sont de l'ordre de 5 à 10 jours pour les radars fixes et de deux semaines à un mois pour les radars mobiles. **Dušek & Traxler (2026)** exploitent les données administratives d'un système de radars automatiques comportant deux sources de variation quasi expérimentale du délai entre l'infraction et l'envoi de l'avis : des retards de traitement allant jusqu'à trois mois au lancement du système, puis un protocole affectant aléatoirement les avis à un traitement rapide ou différé. Ils établissent deux résultats. D'abord, la réception d'un procès-verbal entraîne une **baisse forte, immédiate et persistante des excès de vitesse**. Ensuite, ils ne détectent **aucun effet différentiel robuste de la célérité ou du délai sur le comportement de vitesse ultérieur**, alors même que le délai dégrade nettement le paiement dans les temps. Cette étude ne renseigne en revanche pas la dissuasion générale exercée sur l'ensemble des conducteurs, qui est pourtant le canal principal invoqué en faveur du contrôle automatisé.

4.2 Coûts économiques et sociétaux

Les **coûts directs** comprennent l'installation (ordre de grandeur de 90 000 à 165 000 € par radar fixe selon le type d'équipement, issu de la presse spécialisée¹⁷), la maintenance et le renouvellement (les coûts de maintenance étant estimés autour de 47 M€ en 2023 par la Cour des comptes), auxquels s'ajoute le vandalisme (§2.6). Les **coûts indirects et d'opportunité** sont également discutés : les usagers supportent amendes, pertes de points, temps administratif et effets possibles sur l'assurance ; et les ressources allouées au contrôle automatisé pourraient, en **coût d'opportunité**, financer d'autres mesures de sécurité (infrastructures, aménagements, éducation, ou encore santé publique en général). Par exemple, une étude britannique (Claxton et al., 2015) établit qu'un million de livres sterling

investi dans la santé permet de sauver 5 vies et de gagner 75 années de vie en bonne santé, cumulées sur plusieurs patients évidemment. Le bilan sociétal net dépend donc de la valeur attribuée aux vies épargnées au regard de ces coûts. Des analyses coût-bénéfice ont été publiées pour d'autres programmes, notamment l'évaluation quadriennale du programme national britannique de *safety cameras* (Gains et al., 2005), qui rapporte un rapport bénéfices sur coûts de l'ordre de 2,7 pour 1, établi sur les seules collisions corporelles. Pour le programme français, la Cour des comptes indique elle-même avoir « fait le choix de ne pas évaluer le coût/avantage de la politique, en raison de l'insuffisance des données disponibles » (2021, p. 14) ; nous n'avons identifié aucun bilan coût-bénéfice actualisé et public sur la période récente.

Une limite propre au caractère **automatisé** du contrôle mérite enfin d'être signalée. Parce qu'il sanctionne à distance, par un avis adressé au titulaire de la carte grise, et non par interception, il peut être neutralisé par la fraude à l'immatriculation. Le contrôle sans interception a ainsi possiblement favorisé l'essor des « **garages fantômes** », des sociétés écrans qui immatriculent des véhicules pour le compte de conducteurs échappant de la sorte à leurs procès-verbaux. Une note de la Direction nationale de la police judiciaire (septembre 2025) évalue à plus de **250 000 véhicules par an** le nombre concerné, et la Cour des comptes estime que près d'un million de véhicules circuleraient aujourd'hui hors de portée du contrôle ; ces véhicules ne reçoivent aucun procès-verbal. Le phénomène a été facilité par la dématérialisation de l'immatriculation (accès direct au SIV ouvert à quelque 34 000 professionnels). Cette fuite affaiblit à la fois le rendement et l'équité du dispositif.¹¹

Ordres de grandeur sur la période étudiée (2019-2024). À titre de simple rappel factuel : l'installation des 1 521 radars fixes mis en service entre 2019 et 2024 représente, sur la base de cet ordre de grandeur de 90 000 à 165 000 € l'unité, un investissement de 140 à 250 M€. La maintenance de l'ensemble du parc, environ 47 M€ par an (Cour des comptes), avoisine 280 M€ cumulés sur la période. Les amendes issues des radars ont rapporté de l'ordre de 4,5 à 5 Md€ cumulés sur 2019-2024 (documents budgétaires « jaune » CCSR) ; ces deux derniers chiffres portent sur le parc entier (environ 4 000 dispositifs), et non sur les seuls radars récents, et les amendes constituent une recette publique, non un coût (voir §2.4 sur leur affectation).

4.3 Vie privée et surveillance

L'usage massif du contrôle automatisé soulève des questions de **vie privée et de surveillance**. La lecture automatisée des plaques d'immatriculation (LAPI) permet une collecte systématique de données de déplacement, traitées par l'ANTAI et conservées pour

des durées variables, ce qui pose des enjeux de proportionnalité et de protection des données (conformité RGPD en Europe). Des polémiques récurrentes portent sur le risque de détournement ou de croisement de ces données avec d'autres fichiers administratifs, et le dispositif est régulièrement accusé de participer à une logique de « surveillance de masse », critique avivée lors des périodes de contestation sociale (le vandalisme des « gilets jaunes », §2.6, a pris une dimension symbolique de rejet du contrôle). S'y ajoutent des débats sur l'**opacité** du système (absence de visibilité en temps réel pour le conducteur, délais de notification, difficulté à contester efficacement une infraction) et, à l'international, sur des systèmes plus intrusifs intégrant de la vidéosurveillance.

Le caractère automatisé du contrôle expose enfin à un risque d'**usurpation d'identité** routière : la reproduction frauduleuse de la plaque d'un tiers (« doublette ») fait recevoir à la victime des procès-verbaux pour des infractions qu'elle n'a pas commises. La charge de la preuve pèse alors sur elle : elle doit déposer plainte, puis contester **chaque** avis dans les délais impartis en réunissant des justificatifs d'absence (billets de train, reçus de péage, attestations), l'administration ne résolvant pas le problème à la source. Ces démarches, répétitives et longues, illustrent la difficulté du dispositif à réparer efficacement le préjudice des personnes lésées.¹²

4.4 Contrôler tous pour sanctionner quelques-uns : la question de la proportionnalité

Au-delà des enjeux de protection des données (§4.3), le contrôle automatisé soulève une question de **philosophie du droit** : est-il légitime de soumettre l'ensemble des usagers à une mesure systématique de leur comportement afin de sanctionner la fraction, très minoritaire à l'endroit contrôlé, des passages en infraction ? Un radar fixe classique mesure la vitesse de tous les véhicules qui le franchissent mais n'enregistre que les passages en infraction ; il ne relève donc pas, au sens strict, d'une collecte généralisée et indifférenciée de données visant l'ensemble des usagers, référence à l'aune de laquelle la Cour de justice de l'Union européenne apprécie les dispositifs de surveillance de masse¹³. Les radars tronçon, qui lisent la plaque de chaque véhicule à l'entrée et à la sortie de la zone contrôlée et sont tenus d'en effacer les passages conformes sous 24 heures, se rapprochent davantage de cette limite.

Deux traditions s'opposent sur la légitimité d'un tel contrôle. La première, utilitariste, y voit l'aboutissement du programme de Beccaria (1764) : une sanction modérée mais certaine, prévisible et impersonnelle dissuade mieux qu'une sanction sévère mais aléatoire, et l'automate garantit une **égalité devant la loi** que le contrôle humain, sélectif par nature,

n'assure pas. Le radar annoncé par panneau fonctionne alors comme le panoptique de Bentham analysé par Foucault (1975) : il produit la conformité par la seule conscience d'une observation possible, sans qu'il soit besoin de punir beaucoup.

La seconde tradition, libérale, objecte qu'un contrôle systématique et sans soupçon préalable inverse la logique classique du droit pénal : chaque usager y est traité en contrevenant potentiel, le constat d'infraction est délégué à une machine et la sanction s'applique par amende forfaitaire sans intervention d'un juge, sauf contestation ; l'automate applique en outre la règle sans le correctif d'**équité** que la tradition aristotélicienne confie précisément au juge pour tempérer la généralité de la loi (Aristote, *Éthique à Nicomaque*, V). Elle souligne enfin un problème de **vérifiabilité** : les garanties entourant les données (effacement, durées de conservation, sécurité) sont posées par les textes, mais l'usager n'a aucun moyen propre d'en constater le respect, et aucune certification indépendante régulière n'est publiée. Le principal contrôle connu illustre la portée de l'objection : en 2018, des vérifications sur place de la CNIL ont constaté que les plaques des véhicules en règle relevées par les radars tronçon étaient conservées plus de 13 mois au lieu des 24 heures réglementaires, ce qui a valu au ministère de l'Intérieur une mise en demeure publique, suivie d'un retour à la conformité constaté en 2021¹⁵. L'épisode nourrit les deux lectures : le mécanisme de supervision existe et a fini par documenter puis corriger l'écart (« qui gardera les gardiens ? », demandait Juvénal ; ici, une autorité indépendante) ; mais la garantie d'effacement avait été méconnue pendant des années sans que rien ne le signale publiquement, et sa vérification demeure ponctuelle, tributaire d'un contrôle sur place.

Une partie du débat tient d'ailleurs à la définition même de l'objet sanctionné : quelle est la « juste » vitesse ? Le droit distingue l'**excès de vitesse**, dépassement d'un seuil fixé à l'avance, seule chose qu'un automate sache constater, et la **vitesse excessive**, inadaptée aux circonstances, que l'article R. 413-17 du code de la route réprime y compris en deçà des limites : rouler à la vitesse autorisée peut être dangereux sous la pluie ou aux abords d'une école, et les bilans officiels de l'accidentalité retiennent du reste comme facteur la « vitesse excessive ou inadaptée », catégorie plus large que le seul dépassement de la limite. Le seuil uniforme n'est donc qu'une approximation du danger réel, fixée de loin et à l'avance ; l'agent qui intercepte peut moduler son appréciation d'après la situation qu'il observe, information dont l'autorité centrale ne dispose pas, tandis que l'automate applique l'approximation telle quelle, verbalisant le dépassement formel même bénin et ignorant la vitesse inadaptée mais légale. Ce que le contrôle automatisé gagne en uniformité et en prévisibilité, il le perd en adéquation au contexte ; cet arbitrage recoupe le correctif d'équité évoqué plus haut et relève, lui aussi, du jugement politique.

Le droit positif porte la trace de ces tensions : la présomption de responsabilité du titulaire de la carte grise n'a été admise par le Conseil constitutionnel puis par la Cour européenne des droits de l'homme qu'entourée de garanties, au prix d'aménagements reconnus de la présomption d'innocence et du droit de ne pas s'auto-incriminer¹⁴. Il arbitre ordinairement ce type de conflit par un **contrôle de proportionnalité** : une atteinte aux droits n'est admise que si elle est adaptée (apte à atteindre son but), nécessaire (aucun moyen moins intrusif n'est aussi efficace) et proportionnée au sens strict (les bénéfices l'emportent sur les coûts). Les deux premières branches de ce test sont des questions empiriques, et c'est ici que la philosophie du droit rejoint l'objet de la présente étude : un avant/après non corrigé, qui attribue aux radars une baisse d'environ 30 % des accidents, et un état des données où cette baisse brute ne mesure pas l'effet du radar sans que l'effet localisé sur la fréquence soit pour autant identifié, son signe n'étant pas déterminé (§8, §10), n'alimentent pas la balance avec les mêmes poids. Symétriquement, ce résultat localisé ne clôt pas le débat en sens inverse : l'essentiel de l'effet attendu d'un contrôle systématique relève de la dissuasion générale exercée sur l'ensemble du réseau, la « certitude » beccarienne, qui échappe par construction au périmètre de l'étude (§1, §10).

Que conclure, alors, tant que l'ampleur du bénéfice n'est pas établie ? Le juge, en pratique, n'a pas tranché sur preuves : le contrôle de constitutionnalité s'en tient à l'absence de disproportion manifeste et ne subordonne pas la validation d'un dispositif répressif à la démonstration de son efficacité ; la Cour européenne des droits de l'homme, qui laisse aux États une large marge d'appréciation, leur impose dans le même temps une obligation positive de protéger la vie, y compris par une police effective de la circulation routière¹⁶. Chaque tradition tire ainsi de l'incertitude une conclusion opposée : pour l'une, le doute sur l'ampleur du bénéfice ne justifie pas de renoncer à une mesure peu intrusive susceptible d'épargner des vies ; pour l'autre, une sujétion imposée à tous requiert une preuve positive d'efficacité, faute de quoi le doute doit profiter à la liberté (*in dubio pro libertate*).

5. Données et périmètre

ÉLÉMENT	DÉTAIL
Source accidents	Fichiers BAAC (ONISR / data.gouv.fr)
Période retenue	2017 à 2024 (2017-2018 = années de référence)
Accidents géolocalisés (métropole)	415 623
Radars fixes (métropole)	3 167, dont 1 521 datés 2019-2024 (dont 1 030 remplacent un radar présent au même endroit en octobre 2018)
Source radars	Fichier des radars automatiques (data.gouv.fr) : état actuel, et extraction antérieure (état interne du 5 septembre 2018) pour la nature du traitement
Source trafic	SDES, <i>Bilan de la circulation</i> (contrôle de robustesse)

Fichier des radars. La date de mise en service portée par le fichier des radars est celle de l'équipement en place, non celle du site : lorsqu'une cabine est remplacée par une tourelle, la ligne prend la date de la tourelle et la cabine disparaît du fichier. Pour mesurer la part de ces remplacements, nous avons rapproché le fichier actuel d'une extraction du même fichier antérieure aux dégradations de l'hiver 2018-2019 (état interne du 5 septembre 2018, 3 287 radars en métropole). Le rapprochement est fiable : 96 % des 1 646 radars actuels datés d'avant 2019 s'y retrouvent à moins de 100 m (94 % à moins de 50 m, et pour ceux-là la date de mise en service est la même à un an près dans 99,8 % des cas). Appliqué aux 1 521 radars datés 2019-2024, il donne **1 030 sites rééquipés** (68 % ; un radar présent à moins de 100 m en octobre 2018 et absent du fichier actuel), 19 cohabitations (le radar de 2018 est toujours listé à côté du nouveau) et **472 sites nouveaux** (31 %). La part de rééquipements vaut 88 % pour le millésime 2019 et 91 % pour 2020, puis 58 %, 43 %, 46 % et 46 % de 2021 à 2024 ; dans 99 % des cas, le radar remplacé avait été installé avant 2017, le plus souvent entre 2005 et 2012, et c'était un radar fixe de vitesse (une cabine) dans 913 cas sur 1 030, le radar neuf étant une tourelle dans 905 cas. Pour deux tiers du groupe traité, la période « avant » de l'étude est donc une période déjà sous radar, et l'événement daté est un renouvellement d'équipement, avec le cas échéant une interruption de service que le fichier ne permet pas de mesurer. Cette variable de statut n'entre dans aucune estimation publiée ; elle décrit le traitement (§7.2, §10) et sert de diagnostic de spécification (annexe 13.6, section E).

Pourquoi 2017-2024 ? La géolocalisation GPS des accidents, indispensable pour relier un accident à un radar, est trop lacunaire avant 2017 : la part des accidents dont les coordonnées sont renseignées vaut 14 % en 2005, 37 % en 2014 et 69 % en 2016, contre 87 % en 2017 et 97 % en 2018. Une extension à 2005-2016 a été envisagée puis écartée : les données anciennes ne permettent pas de localiser les accidents assez précisément pour un rayon de 500 m.

Pourquoi la France métropolitaine seulement ? Les départements et territoires d'outre-mer ont été exclus pour trois raisons. D'abord la **comparabilité** : réseaux routiers, climat, volumes de trafic et accidentalité y diffèrent fortement de la métropole, et les y mêler introduirait de l'hétérogénéité. Ensuite la **disponibilité des données** : l'open data ne couvre certains territoires ultramarins que depuis 2019. Enfin le **volume** : les effectifs ultramarins sont faibles et le déploiement de radars y suit une logique différente. Le filtre géographique appliqué (latitude 41-52, longitude -5,5 à 10) isole proprement la métropole et la Corse, et écarte les points ultramarins.

Indicateurs de gravité. Les *tués* (grav=2) reposent sur une définition constante sur toute la période : est tuée toute personne qui décède sur le coup ou dans les trente jours suivant l'accident, définition en vigueur depuis le 1er janvier 2005 (loi du 9 août 2004, arrêté du 27 mars 2007, conforme à la décision européenne 93/704/CE créant la base CARE) et inchangée depuis (ONISR, *Description des bases de données annuelles*, octobre 2025). Aucune rupture de série ne concerne donc la mortalité sur 2017-2024, et l'indicateur est labellisé par l'autorité de la statistique publique. Les *blessés hospitalisés* (grav=3), en revanche, ont fait l'objet d'un changement de saisie en 2018 (indicateur non labellisé depuis 2019) : leur nombre chute de ~25 400 en 2017 à ~20 800 en 2018, puis ~18 000 en moyenne de 2019 à 2024, sans amélioration réelle. Toute analyse incluant les blessés graves est donc restreinte à **2018-2024**, fenêtre où la saisie est homogène. Enfin, depuis les données 2021, les usagers en fuite sont ajoutés à la base avec des informations lacunaires (sexe, âge, voire gravité) : ne pouvant per se être comptés comme tués, et leur gravité étant le plus souvent non renseignée, ils n'affectent pas les indicateurs utilisés ici.

6. Aperçu descriptif des données

Avant toute analyse statistique approfondie, il est utile d'examiner les données brutes, telles qu'elles apparaissent. Le tableau suivant récapitule l'accidentalité **des accidents géolocalisés retenus** (France métropolitaine) sur la période. Ces comptes ne sont pas les totaux nationaux : ils excluent les accidents non géolocalisés (13 % en 2017, 3 % en 2018, marginal ensuite), d'où un écart avec les bilans ONISR sur les deux premières années. Leur total sur huit ans fait les 415 623 accidents de l'étude.

ANNÉE	ACCIDENTS	TUÉS	BLESSÉS GRAVES	BLESSÉS LÉGERS	ACCIDENTS MORTELS	TUÉS POUR 100 ACCIDENTS
2017	51 866	3 280	25 448*	39 360*	3 067	6,32
2018	53 997	3 178	20 759	46 956	2 984	5,89
2019	56 023	3 247	19 612	50 883	3 052	5,80
2020	45 121	2 541	15 600	40 236	2 403	5,63
2021	53 540	2 944	17 852	49 205	2 771	5,50
2022	52 380	3 267	18 024	47 406	3 077	6,24
2023	51 641	3 167	18 051	46 623	2 984	6,13
2024	51 055	3 193	17 831	46 700	3 011	6,25

* La répartition graves/légers de 2017 n'est pas comparable aux années suivantes (changement de saisie en 2018) ; le total des blessés, lui, reste cohérent.

On notera que, dans nos données, 2017 est l'année la plus meurtrière de la période (3 280 tués), un élément qui jouera un rôle dans l'interprétation des résultats sur la mortalité (§8).

Mortalité à proximité des radars. Le nombre de tués survenus à moins de 500 m d'un radar (tous millésimes confondus) est le suivant, en comptage brut puis après correction de la couverture GPS appliquée dans tous les modèles (§7.3) :

ANNÉE	2017	2018	2019	2020	2021	2022	2023	2024
Tués à moins de 500 m d'un radar (brut)	157	123	163	135	154	142	126	143
Après correction de la couverture GPS	180	127	163	135	154	142	126	143

Une fois cette correction appliquée (moyenne 146,3 par an), la série ne présente pas de tendance nette hors son point de départ 2017, plus élevé que le comptage brut ne le laissait voir. Cette série est descriptive : elle porte sur l'ensemble du parc, et non sur le seul groupe de radars récents dont les modèles du §8 estiment l'effet.

Un constat brut, et ce qu'il ne prouve pas. Si l'on se contente d'un avant/après descriptif autour des radars installés depuis 2019, on observe une baisse apparente. Dans un rayon de 500 m, hors année de mise en service (année de transition, exclue des comptes comme de l'exposition), **5 869 accidents** ont été recensés *avant* installation et **2 679 après**. Ces comptes bruts ne sont **pas directement comparables** : la fenêtre « avant » et la fenêtre « après » n'ont pas la même durée, puisque les radars ont été installés à des dates échelonnées entre 2019 et 2024 et que les données s'arrêtent fin 2024. En moyenne, chaque zone est observée plus longtemps avant qu'après : **4 247 années-zone avant, 2 767 après**. Une fois rapportés à cette **exposition**, les taux annuels par zone font apparaître des baisses d'environ **-30 % pour les accidents** (1,38 → 0,97 accident par zone et par an, 1 002 zones, 4 247/2 767 années-zone) et **-31 % pour les tués** (298 → 134 en comptes bruts, même exposition). Pour les **blessés graves**, la fenêtre homogène 2018-2024 (§5) impose un univers légèrement différent (948 zones, 3 081/2 607 années-zone) : la baisse y est de **-34 %** (1 385 → 771). Sur la fenêtre longue 2017-2024, la baisse des blessés graves est de **-37 %**, mais elle inclut 2017, année que le changement de saisie rend non comparable aux suivantes (§5) : c'est le **-34 %** qui fait référence. La convention d'univers du constat descriptif (zones ayant enregistré au moins un accident dans la fenêtre considérée) est sans effet sur les comptages, les zones sans accident contribuant zéro ; elle est détaillée à l'annexe 13.4. Ces chiffres descriptifs, non corrigés du retour à la moyenne, suggèrent *a priori* un effet protecteur des radars. C'est précisément cet *a priori* que la suite de l'étude met à l'épreuve : cette baisse de 30 % ne mesure pas l'effet du radar (annexe 13.7), sans que ces données permettent pour autant de déterminer le signe de l'effet local sur la fréquence (§8). La démarche statistique (§7) repart de ce constat brut.

7. Choix méthodologiques

7.1 Le rayon d'influence : pourquoi 500 m et 1 km ?

Un radar fixe ne modifie le comportement des conducteurs que là où ils le perçoivent : pour l'essentiel, la portion comprise entre le panneau d'annonce et la cabine, plus le voisinage immédiat. L'étendue de cette zone d'influence est donc bornée par la **doctrine d'implantation officielle**.

L'arrêté du 4 mars 2013 (§12.3) prévoit que la zone contrôlée peut être signalée par un panneau d'annonce (modèles SR3a ou SR3b), dans le cadre de l'instruction interministérielle sur la signalisation routière ; la doctrine d'implantation qui l'accompagne situe ce panneau à une distance de la cabine qui croît avec la vitesse maximale autorisée (VMA), selon les ordres de grandeur suivants :

VMA	TYPE DE VOIE	DISTANCE PANNEAU → CABINE
130 km/h	Autoroute	~600 m (± 50 m)
110 km/h	Voie express / rapide	~500 m
80-90 km/h	Hors agglomération	~400 m
70 km/h	Zones intermédiaires	~300 m
50 km/h	Agglomération	pas de distance standardisée

La zone de ralentissement effectif d'un radar fixe s'étend ainsi typiquement sur **300 à 600 mètres en amont**. C'est ce qui motive le choix des deux rayons :

- **500 m (rayon principal)** : couvre la distance panneau → cabine pour la plupart des types de routes. Les accidents étant enregistrés à leur lieu de survenue (en amont, là où l'on freine, ou à hauteur de la cabine), 500 m capture la zone où le radar peut influencer.
- **1 km (contrôle de robustesse)** : zone plus large, captant une éventuelle influence étendue ou des effets de débordement, au prix d'inclure des accidents moins plausiblement liés au radar (dilution).

Au-delà de ~1 km, les accidents sont essentiellement sans rapport avec le radar (aucune signalisation, aucune influence comportementale à cette distance) : c'est pourquoi un rayon de 2 km a été écarté, il n'ajoute que du bruit, sans information identifiante.

La présence du panneau n'est **pas** une condition de validité du procès-verbal : l'arrêté du 4 mars 2013 dispose que la signalisation « peut être effectuée » au moyen du panneau SR3, sans en faire une condition de la constatation de l'infraction. Les distances ci-dessus relèvent donc d'une **doctrine technique** d'implantation que l'État s'impose (éviter les freinages brusques), non d'une obligation dont la méconnaissance entacherait le procès-verbal. Sur certains grands axes accidentogènes, la doctrine a en outre évolué vers des **zones de contrôle** signalées globalement (« contrôles radars fréquents sur 25 km »), où les panneaux individuels avant chaque cabine sont supprimés : sur ces axes, l'hypothèse d'une influence localisée à 500 m-1 km est plus faible.

7.2 Le groupe témoin : zones à radars anciens

Un DiD compare l'évolution des zones traitées à celle d'un groupe témoin. Deux options se présentaient :

Option A (retenue) : témoins = abords des radars installés avant 2019. Ces zones ressemblent fortement aux zones traitées : un emplacement reçoit un radar parce que c'est un point dangereux (trafic dense, accidents fréquents, route rapide). Radars anciens et récents sont donc sur des sites de **même nature**, ce qui est le fondement d'un bon DiD : on compare des choses comparables, réduisant le risque que la différence vienne d'ailleurs. Cette similitude porte sur la logique d'implantation, non sur la technologie ni sur le niveau de base : les deux groupes diffèrent par la composition (87 % de tourelles côté traité ; cabines fixes, radars de feu rouge, discriminants, tronçon et passage à niveau côté témoin, annexe 13.6, section D) et les zones traitées comptent environ trois fois moins d'accidents que les témoins (§11). Un contrôle restreignant les traités aux tourelles et les témoins aux cabines fixes et discriminants donne +0,6 % ($p = 0,91$). Pour deux tiers du groupe traité, la similitude est même une identité de site : 1 030 des 1 521 radars datés 2019-2024 ont pris la place d'un radar présent au même endroit en octobre 2018 (§5). Le contraste estimé n'oppose donc pas des sites nouvellement équipés à des sites équipés de longue date, mais des sites dont l'équipement a été renouvelé ou, pour un tiers, créé, à des sites dont l'équipement n'a pas changé. Limite de ce choix : sur les zones à radar ancien, l'effet du radar a déjà agi avant 2019. Un témoin n'a pas besoin d'être « sans radar », et l'absence de *changement* de radar sur la période est vérifiée pour 1 594 des 1 646 radars témoins ; les 52 restants (46 mis en service en 2017, 6 en 2018) changent de statut à l'intérieur de la fenêtre d'observation, 36 d'entre eux sont le radar le plus proche d'au moins un accident survenu à moins de 500 m et constituent donc une zone témoin de l'estimation, et leur exclusion laisse l'estimation inchangée (+3,6 % au lieu de +3,4 %, annexe 13.6). Mais cette absence de changement est une

condition nécessaire, non la condition d'identification : celle-ci est le parallélisme des tendances, qui suppose en outre que l'effet propre des radars témoins soit stabilisé sur la période (§11).

Option B (écartée) : témoins = zones tirées au hasard loin de tout radar (> 2 km). Intuitivement plus proches de l'idée de « contrôle », mais piégées par un **biais de sélection** : on ne place pas un radar n'importe où. « Loin de tout radar » désigne souvent des routes plus calmes, moins accidentogènes par nature. On comparerait alors des zones très différentes et l'effet mesuré serait pollué par cette différence de fond.

L'Option B et une variante « sites bientôt équipés » (radars récents observés *avant* leur propre installation) ont été explorées en phase préparatoire sans être conduites sous une forme publiable, et aucun résultat de notre étude ne s'y appuie ; la seconde (« sites bientôt équipés ») est toutefois reprise, sous une forme contrôlée, par la variante A de l'estimateur de Callaway-Sant'Anna (annexe 13.5 : -3,2 %, non significatif). Le programme de reproduction publié permet à un lecteur d'estimer l'une et l'autre.

7.3 La démarche en quatre étapes

1. Comparaison avant/après naïve. À exposition comparable, elle suggère une baisse de **-30 %** du nombre d'accidents (§6).

2. Détection du retour à la moyenne. Un DiD sans correction de pente afficherait lui aussi une baisse trompeuse, d'environ -29 % ($p < 0,001$, annexe 13.1). Mais une étude d'événement révèle que la trajectoire d'accidentalité des zones traitées est déjà en mouvement **avant** l'installation. Examinée à l'intérieur des cohortes, c'est-à-dire à composition constante sur les cinq années qui précèdent la pose, sur les cohortes 2022-2024 (414 zones traitées sur 1 002, seules à disposer de cinq à sept années antérieures à la pose), cette trajectoire est **croissante** et culmine l'année qui précède immédiatement la pose (annexes 13.1 et 13.7) : c'est la signature d'une sélection réactive sur un pic. L'étude d'événement en coupe poolée, qui mélange des cohortes d'installation différentes, donne une image différente et partiellement trompeuse de cette trajectoire (annexe 13.6, section B). Les radars étant posés en réaction à un pic d'accidents, ces baisses apparentes sont pour l'essentiel un artefact de sélection réactive et de composition, et non un effet causal : la décomposition de l'annexe 13.4 montre que l'écart au chiffre corrigé provient pour environ deux tiers du passage à une pondération site à site et pour un tiers de la tendance du groupe témoin, la part strictement imputable au retour à la moyenne n'étant identifiée que par le terme de pente du test de cassure, que rien dans les données ne distingue d'une autre

tendance propre aux sites traités. En résumé, la baisse observée après la pose tient d'abord à la façon dont les sites ont été choisis et dont les chiffres ont été comptés, et l'on ne peut pas dire quelle part exacte revient au simple retour à la normale après un pic ; ce que l'on peut dire, c'est qu'elle ne mesure pas l'effet du radar.

3. Correction. Le **test de cassure (régression de Poisson)** sépare une tendance continue propre aux zones traitées, dont le retour à la moyenne fait partie, d'une rupture nette à l'installation ; c'est lui qui porte l'estimation. L'**Empirical Bayes (formule de Hauer)**, calibré ici sur les sites traités eux-mêmes, ne constitue pas une seconde correction indépendante : il fournit une borne basse et un diagnostic de cohérence (§3.7, annexe 13.4).

4. Robustesse. Deux rayons ; deux modèles de tendance (témoin vs trafic national réel, creux Covid 2020 inclus ; résultats en annexe 13.4) ; correction d'exposition pour la couverture GPS ; erreurs-types regroupées par site. Deux anomalies de données ont été détectées et corrigées : sur l'année 2022 (identifiant d'accident mal converti, gravité non jointe) et sur l'identification des cabines (701 radars sans numéro dans le fichier source, dont 101 installés 2019-2024, qui échappaient à la constitution des zones ; un identifiant de ligne, propre à chaque cabine, les réintègre).

Précision sur le terme d'exposition (offset). Dans toutes les régressions, chaque zone-année est pondérée par la part des accidents géolocalisés de l'année concernée (87 % en 2017, 97 % en 2018, quasi 100 % ensuite) : c'est ce facteur de **couverture GPS**, noté c_t , qui sert d'offset $\log(c_t)$ dans les modèles (annexe 13.3). Ces deux valeurs sont la part nationale, outre-mer compris, des accidents dont les coordonnées sont renseignées (87,3 % et 96,8 % avant arrondi, étape 15 du programme de reproduction) ; mesurée sur la seule métropole avec le critère de sélection des accidents retenus, la couverture vaut 88,5 % en 2017 et 96,8 % en 2018 (étape 14 du programme de reproduction). La différence est sans effet sur les régressions, l'offset ne dépendant que de l'année et se confondant avec les effets fixes d'année ; elle déplacerait d'au plus trois unités la série des tués corrigée du §6. Il faut d'ailleurs le dire nettement. Dans les régressions, qui comportent toutes des effets fixes d'année, ce terme fait double emploi avec eux et ne déplace pas l'estimation : le retirer laisse le résultat inchangé au dix-millième près (étape 17 du programme de reproduction). Il n'agit que dans l'Empirical Bayes, dont l'attendu dépend de la somme des couvertures, et dans les séries descriptives du §6, où sans lui la moindre complétude des données 2017-2018 serait comptée à tort comme un surcroît d'accidents après installation. Ce n'est **pas** un volume de trafic. Le trafic national (SDS) n'intervient, lui, qu'en contrôle de robustesse, comme modèle de tendance alternatif au groupe témoin.

8. Résultats consolidés

COMMENT LIRE CE TABLEAU

Chaque ligne donne l'effet du radar estimé à l'installation (la rupture), après correction du retour à la moyenne : un effet **néгатif** est une baisse, un effet **positif** une hausse. Le **seuil détectable** est la plus petite baisse repérable de façon fiable (puissance de 80 %) ; en deçà, un effet réel resterait invisible. Rayon : distance autour du radar sur laquelle on mesure l'influence (500 m ou 1 km). IC 95 % : fourchette des effets compatibles avec les données (la procédure qui la construit encadre la vraie valeur dans 95 % des cas). p : probabilité d'observer un écart au moins aussi grand que celui mesuré si le radar n'avait aucun effet ; un p petit (sous 0,05) signale un résultat difficilement attribuable au seul hasard. KSI = Killed or Serious Injured (tués ou blessés graves). La colonne Lecture porte deux verdicts qu'il ne faut pas confondre : **aucun effet détectable** signifie que les données ne montrent pas d'effet et n'ont pas la sensibilité d'en exclure un, la limite étant le nombre d'événements ; **effet non identifié** signifie que la méthode elle-même ne restitue pas zéro quand elle le devrait (voir la réserve énoncée plus bas et l'annexe 13.7), de sorte que ni le sens ni l'ampleur de l'effet ne sont établis.

Note de nomenclature. Ce que le rapport appelle « blessés graves » est, dans les fichiers BAAC, le **blessé hospitalisé** (grav = 3). Ce n'est pas le blessé grave au sens MAIS3+ retenu par l'ONISR pour ses estimations de référence, et l'indicateur n'est plus labellisé depuis 2019 (§5). Il s'agit donc d'un proxy imparfait, et c'est la raison pour laquelle l'analyse de gravité est restreinte à la fenêtre de saisie homogène 2018-2024.

INDICATEUR	RAYON	EFFET	P	IC 95 %	SEUIL DÉTECTABLE	LECTURE
Fréquence (tous accidents)	500 m	+3,4 %	0,52	[−7 % ; +15 %]	~14 %	effet non identifié
Fréquence	1 km	+4,9 %	0,27	[−4 % ; +14 %]	~11 %	effet non identifié
Tués (2017-2024)	500 m	−17 %	0,38	[−45 % ; +25 %]	~44 %	aucun effet détectable (voir la divergence EB ci-dessous)
Tués (2017-2024)	1 km	−7,5 %	0,59	[−30 % ; +23 %]	~33 %	aucun effet détectable (voir la divergence EB ci-dessous)
Blessés graves (2018-2024)	500 m	−1,5 %	0,89	[−20 % ; +21 %]	~25 %	aucun effet détectable
Blessés graves (2018-2024)	1 km	+8,6 %	0,26	[−6 % ; +25 %]	~18 %	aucun effet détectable
KSI (tués + graves, 2018-2024)	500 m	−2,9 %	0,77	[−20 % ; +18 %]	~24 %	aucun effet détectable
KSI (2018-2024)	1 km	+6,7 %	0,36	[−7 % ; +22 %]	~18 %	aucun effet détectable

Sur la fréquence, les seuils correspondants sont de 11 à 14 % à la baisse et de 13 à 16 % à la hausse ; la colonne « Seuil détectable » du tableau reporte la borne à la baisse.

Les estimations Empirical Bayes (borne basse et diagnostic, non une seconde correction indépendante, §3.7) sont cohérentes sur la fréquence : -3,4 % (500 m) et -2,0 % (1 km), non significatifs ; une variante de robustesse prenant le trafic national comme tendance donne -7,5 % / -5,5 % (annexe 13.4). Ces variantes font varier le contrefactuel à forme de tendance constante, et leurs résultats restent proches ; l'annexe 13.6 montre que faire varier la forme de la tendance les déplace bien davantage. Sur les tués, l'EB estime en revanche une baisse (-30 % / -24 %, IC excluant zéro) que le test de cassure ne retrouve pas : cette divergence entre le test de cassure et l'Empirical Bayes, discutée ci-dessous et en annexe 13.4, est le principal signal d'instabilité des résultats de mortalité. Un contrôle d'adoption échelonnée (estimateur de type Callaway-Sant'Anna, annexe 13.5) donne par ailleurs, sur la fréquence, des estimations du même ordre, également non significatives.

Interprétation Aucun des huit tests de cassure n'est statistiquement significatif : les valeurs de p s'échelonnent de 0,26 à 0,89 et aucune n'approche le seuil de 5 %. Toute correction de multiplicité, qui ne peut qu'augmenter ces valeurs, laisserait la lecture inchangée. Six des huit estimations Empirical Bayes sur la fréquence et les tués excluent en revanche zéro. Ce sont les quatre estimations sur les tués (-30,1 % à 500 m et -24,3 % à 1 km avec la tendance mesurée sur les tués, $p < 0,001$; -21,7 % et -11,9 % avec la tendance mesurée sur les accidents, $p = 0,006$ et $p = 0,042$) et les deux estimations de fréquence de la variante à tendance trafic (-7,5 %, $p = 0,003$; -5,5 %, $p < 0,001$). Les deux estimations de fréquence à tendance témoin, elles, ne sont pas significatives (voir la note ci-dessus et l'annexe 13.4). Sur la gravité, une des quatre estimations Empirical Bayes de la fenêtre 2018-2024 exclut zéro (KSI à 500 m, point 1 ci-dessous). Le statut de ces estimations reste celui d'un diagnostic de cohérence et d'une borne basse, non d'une correction indépendante du retour à la moyenne (§3.7). On note qu'avec huit tests indépendants au seuil de 5 %, la probabilité qu'au moins un ressorte « significatif » en l'absence de tout effet serait d'environ une sur trois (les tests étant en réalité corrélés, cette probabilité est un majorant) : une estimation isolée et significative, parmi la centaine et plus publiée dans cette étude (annexes 13.4 à 13.7), ne suffirait donc pas à établir un effet.

Une réserve qui porte sur la fréquence des accidents. L'estimation de +3,4 % suppose une forme donnée au reflux d'accidentalité qui précède l'installation. Cette hypothèse n'est pas vérifiable, et elle n'est pas neutre : d'autres formes, également défendables, déplacent l'estimation de plusieurs dizaines de points (annexe 13.6). Surtout, un test placebo attribue aux radars une fausse date de mise en service, antérieure de deux ou trois ans à la date réelle, et n'exploite que les années précédant l'installation : dans cette fenêtre, l'effet vaut zéro par construction. La spécification retenue ne produit pas ce zéro : elle s'en écarte de 19,2 points en médiane, toujours dans le même sens, significativement dans trois des sept

configurations testées (annexe 13.7). Ce test ne porte que sur les cohortes 2022-2024, seules à disposer d'assez d'années antérieures : il ne dit rien des cohortes 2019-2021, qui fournissent l'essentiel de l'information postérieure à la pose. Cet écart vaut 1,8 fois la demi-largeur de l'intervalle de confiance publié (10,7 points, pour un intervalle large de 21,4 points) et dépasse le seuil de détection de l'estimation.

Conclusion sur la fréquence des accidents. L'étude ne mesure pas d'effet localisé, et ne peut pas non plus affirmer qu'il est nul : le signe de cet effet n'est pas déterminé par ces données. Deux résultats sont en revanche solidement établis. Aucune spécification comportant un terme de tendance ne soutient une baisse de l'ordre de 30 % sur les sites équipés ; seule la comparaison sans terme de tendance la reproduit (−28,9 % à 500 m, −33,4 % à 1 km, annexe 13.6), et c'est aussi celle dont le test placebo établit le plus nettement le biais : le −30 % du §6 ne mesure pas l'effet du radar. Et le même dispositif appliqué à une fausse date d'installation, sans correction de tendance, produit un écart de +8,6 % à +20,4 %, significatif dans six configurations sur sept, la septième (g−3, 500 m, hors 2020) valant +8,6 % avec $p = 0,107$: l'accidentalité d'un site monte jusqu'à l'année qui précède immédiatement la pose, dernière année observée avant l'installation, celle-ci étant exclue comme transitoire ; cela documente directement le pic sur lequel les radars sont implantés.

Sur la gravité et la mortalité, la conclusion est inchangée sous les variantes testées : aucun effet n'est détectable. Ni la spécification retenue ni la variante à pente antérieure seule (annexe 13.6, section A) ne font apparaître d'effet significatif sur les blessés graves, les KSI ou les tués. La réserve de spécification énoncée ci-dessus n'est pas pour autant levée sur ces indicateurs : le test placebo appliqué aux tués, aux blessés graves et aux KSI sur l'ensemble du groupe traité (douze configurations, étape 17 du programme de reproduction) s'écarte de zéro de 26,7 points en médiane, toujours dans le sens d'une baisse, significativement dans quatre configurations sur douze, avec des seuils de détection propres de 28 % à 67 % à la baisse qui ne permettent ni de certifier ni d'exclure un biais de spécification (annexe 13.7 ; détail par sous-groupe à l'annexe 13.6, section E). Les seuils de détection restent de 18 à 25 % à la baisse (22 à 34 % à la hausse) sur la gravité et de 33 à 48 % à la baisse (50 à 94 % à la hausse) sur la mortalité.

Le cas des tués Le test de cassure ne détecte rien (−17 % à 500 m, $p = 0,38$), très en deçà du seuil détectable (−44 %) ; mais l'Empirical Bayes estime une baisse de −30 % [−44 ; −16]. Trois raisons de ne pas conclure pour autant à un effet sur la mortalité :

1. **La divergence entre l'estimation principale et le diagnostic Empirical Bayes** est précisément le signal d'un résultat instable, là où les deux estimateurs concordent sur la fréquence des accidents. Sur la gravité, où l'Empirical Bayes a été estimé sur la fenêtre 2018-2024 (annexe 13.4), l'écart est intermédiaire : à 500 m, l'EB donne -6,3 % sur les blessés graves et -8,8 % sur les KSI, contre -1,5 % et -2,9 % en cassure, et seul le KSI y exclut zéro (IC [-17,0 ; -0,6], $p = 0,036$) ; à 1 km, les deux estimateurs sont de signes opposés (-0,8 % contre +8,6 %, -2,5 % contre +6,7 %), tous non significatifs. Le poids médian w_y vaut 0,22 à 500 m et 0,13 à 1 km : comme sur la fréquence, le contrefactuel repose surtout sur l'historique propre du site, et l'EB reste une borne basse, non une correction du retour à la moyenne (§3.7). Il ne peut donc pas porter une conclusion d'existence sur les KSI, pas plus que sur les tués.
2. **L'estimation EB est sensible à la spécification** : selon la série retenue pour la tendance du groupe témoin, elle passe de -30 % à -22 % (500 m) et de -24 % à -12 % (1 km, borne haute frôlant zéro) (annexe 13.4).
3. **Aucune corroboration sur la gravité.** Le principal canal par lequel un radar peut agir sur la mortalité est la baisse de la vitesse, qui réduit *conjointement* le nombre de tués **et** de blessés graves (relation vitesse-gravité). Or les blessés graves (-1,5 %) et les KSI (-2,9 %), six à sept fois mieux alimentés et soumis à ce même mécanisme, ne montrent aucun effet détectable, à des seuils de 18 à 25 % à la baisse (22 à 34 % à la hausse) : si une baisse de mortalité de l'ampleur suggérée par l'Empirical Bayes était réelle, elle devrait se lire aussi sur ces indicateurs mieux mesurés. Son absence oriente vers un artefact ; sans pour autant l'exclure, la mortalité étant trop rare pour être tranchée à elle seule (seuil ~44 %). Trois réserves tempèrent cet argument : la relation vitesse-gravité, toutes choses égales par ailleurs, est non linéaire (*power model*), de sorte qu'un même effet sur la vitesse n'a pas la même détectabilité sur les tués et sur les blessés graves ; le radar peut agir par d'autres voies que la vitesse moyenne (dispersion des vitesses, composition du trafic, freinages) ; et les KSI contenant les tués, ils ne constituent pas une corroboration pleinement indépendante : seuls les blessés graves le sont.

Un élément de contexte complète cette lecture : 2017, année la plus meurtrière de la période (§6), appartient à la fenêtre « avant » de toutes les zones traitées ; ce point de départ élevé gonfle mécaniquement la baisse apparente sur les tués. La restriction à la fenêtre 2018-2024 réduit d'ailleurs la cassure estimée de -17 % à -11 % (annexe 13.2).

9. Interprétation des résultats

Fréquence des accidents : effet non identifié. C'est l'indicateur le mieux alimenté, et les estimations y sont stables d'une méthode et d'un rayon à l'autre (+3,4 % / +4,9 % en cassure, -3,4 % / -2,0 % en Empirical Bayes, toutes non significatives) : la rupture à l'installation est indistinguable de zéro. Mais cette stabilité est conditionnelle à la forme donnée au reflux antérieur à la pose, et le test placebo (annexe 13.7) établit que la spécification retenue ne restitue pas zéro là où l'effet vaut zéro. On ne peut donc pas conclure ici à l'absence d'effet, seulement à l'impossibilité de le mesurer : le signe de l'effet localisé sur la fréquence n'est pas déterminé par ces données. La contrepartie, établie au §8, est inchangée : aucune spécification comportant un terme de tendance ne soutient une baisse de l'ordre de 30 % sur les sites équipés.

Gravité (blessés graves, KSI) : pas d'effet détectable. Avec ces indicateurs, la puissance suffit à repérer un effet dès ~18-25 %. À ce niveau, aucun effet n'apparaît : ruptures proches de zéro, voire légèrement positives à 1 km, toutes non significatives.

Tués seuls : un indice non concluant. Le test de cassure ne détecte rien (-17 % à 500 m, $p = 0,38$) ; l'Empirical Bayes suggère une baisse, mais sensible aux choix de spécification et **non corroborée** par les blessés graves, six fois plus nombreux et soumis au même mécanisme vitesse → gravité (§8). L'Empirical Bayes étant ici calibré sur la période « avant » des sites traités eux-mêmes (§3.7, annexe 13.4), il ne peut de surcroît pas servir d'argument d'existence d'un effet : il reproduit d'ailleurs presque exactement, sur les tués, l'avant/après non corrigé du §6 (-30,1 % contre -31,0 %). Cette configuration oriente vers le bruit ou un résidu de retour à la moyenne plutôt que vers un effet causal établi ; elle ne permet pas non plus de l'exclure, la puissance étant la plus faible sur cet indicateur.

En nombres absolus : des fourchettes trop larges pour trancher. Sur la fenêtre post-installation observée (2,8 ans par zone en moyenne, rayon 500 m ; calcul détaillé en annexe 13.4), les résultats donnent, selon la méthode et l'intervalle retenus : pour les accidents, entre environ 360 évités et 340 supplémentaires ; pour les blessés graves, entre environ 190 évités et 130 supplémentaires ; pour les tués, entre environ 110 évités et 27 supplémentaires. Ces conversions reprennent les intervalles du test de cassure et, pour les accidents, ceux des variantes Empirical Bayes (annexe 13.4) : sur les accidents, la réserve du §8 s'applique et interdit de lire cette fourchette comme un encadrement de l'effet réel. Les résultats changent de signe, ce qui explique pourquoi l'étude ne publie aucun bilan en « vies sauvées » ni en « victimes supplémentaires » : à l'échelle locale des radars récents, les données ne

permettent pas d'identifier le sens du bilan. On rappelle qu'il s'agit du seul bilan localisé (moins de 500 m des radars, toute année confondue, périmètre où surviennent environ 4,6 % des tués géolocalisés de la période ; environ 2 % si l'on s'en tient aux seuls radars récents), hors tout effet de réseau : la dissuasion générale que le contrôle automatisé peut exercer sur les comportements de conduite partout, y compris loin des cabines, reste en dehors de ce que l'étude mesure (§1, §10).

Ce que « aucun effet détectable » signifie. En statistique, « aucun effet détectable » ne signifie pas « effet nul » : les données suffisent à constater qu'on n'observe pas d'effet, pas à démontrer son absence. Concrètement, les seuils détectables valent environ 11-14 % à la baisse (13-16 % à la hausse) sur la fréquence, 18-25 % (22-34 %) sur la gravité et 33-48 % (50-94 %) sur la mortalité ; ils ne sont pas symétriques, l'échelle du modèle étant logarithmique ; ils décrivent la sensibilité du protocole, non une plage d'effets exclue. Sur la gravité, ce sont les intervalles de confiance du §8 qui bornent, sous la spécification retenue, les effets compatibles avec les données : à 500 m, ils excluent une baisse comme une hausse supérieures à environ 20 % ($[-20 \% ; +21 \%$] sur les blessés graves, $[-20 \% ; +18 \%$] sur les KSI) ; à 1 km, où l'estimation est plus précise, ils excluent une baisse supérieure à 6 à 7 % ($[-6 \% ; +25 \%$] et $[-7 \% ; +22 \%$). Un effet d'au moins 18 à 25 % à la baisse (22 à 34 % à la hausse) aurait eu quatre chances sur cinq d'être détecté ; un effet plus modeste resterait invisible. La forme donnée à la tendance y déplace aussi les estimations (sous la variante à pente antérieure seule, les blessés graves passent de $-1,5 \%$ à $-13,1 \%$ $[-26,7 ; +2,9]$ et les KSI de $-2,9 \%$ à $-13,8 \%$, annexe 13.6, section A), sans qu'aucune variante testée n'y fasse apparaître d'effet significatif : ce qui borne la conclusion sur la gravité reste la puissance, et non l'échec au placebo qui affecte la fréquence, le placebo appliqué à ces indicateurs étant lui-même sans puissance suffisante pour trancher (seuils de 28 % à 67 % à la baisse, annexe 13.7). Sur la seule mortalité, même un effet substantiel (jusqu'à 33-48 % à la baisse, 50-94 % à la hausse) resterait indétectable. Sur la fréquence, en revanche, le seuil d'environ 14 % ne borne pas l'effet réel : il ne le ferait que si la spécification qui le produit était non biaisée, or, au test placebo, elle s'écarte de zéro de 19,2 points en médiane, écart supérieur à ce seuil, et les estimations publiées en annexe 13.6 vont de -16% à $+31 \%$ selon la forme donnée à la tendance. Ces seuils sont calculés *ex post* et ne valent que pour la spécification retenue au §8 ; l'annexe 13.6 montre qu'ils varient avec elle.

Pourquoi une telle différence avec la littérature ? La plupart des évaluations qui corrigent le retour à la moyenne conservent un bénéfice local modéré, de l'ordre de 12 à 22 % (§3.5) ; la présente étude ne le retrouve pas, sans pouvoir établir qu'il est absent (§3.7). Plusieurs propriétés du protocole jouent dans le sens d'une estimation locale plus proche de zéro que

celle du champ, et il est utile de les réunir. Le groupe témoin est fait d'abords de radars anciens : la grandeur estimée est l'effet marginal d'un radar supplémentaire dans un réseau déjà densément équipé, et toute dissuasion commune aux deux groupes sort de la comparaison par différence (§10). Le traitement lui-même n'est pas, pour l'essentiel, une première installation : 1 030 des 1 521 radars datés 2019-2024 (68 %) remplacent un radar présent au même endroit en octobre 2018 (§5). La période « avant » de ces sites était donc déjà sous radar, sous réserve des interruptions de service que le fichier ne permet pas de mesurer, et aucun effet de première arrivée n'y est attendu. La littérature, elle, évalue le plus souvent l'introduction d'un radar sur un site qui n'en avait pas, situation qui ne correspond ici qu'au tiers du groupe traité ; les estimations par sous-groupe ne permettent pas d'isoler l'effet de cette première installation, la spécification y échouant au placebo plus nettement encore que sur l'ensemble (annexe 13.6, section E). L'estimation porte sur l'installation d'une cabine, active ou non : un radar hors service compte comme un radar, ce qui minore l'effet propre d'un dispositif en fonctionnement, dans une proportion que le fichier des radars ne permet pas de mesurer (part du parc en fonctionnement passée d'environ 88 % à 81 % en 2024). Ce canal joue par la sanction et non par la dissuasion visuelle, une tourelle non équipée étant indiscernable d'une tourelle équipée pour le conducteur qui passe (§11.1). Le disque de 500 m est large au regard de la distance à laquelle la littérature situe l'effet (quelques centaines de mètres, et 137 m seulement dans l'évaluation new-yorkaise), et une erreur de position des accidents, non mesurable dans les BAAC, atténuée vers zéro tout effet localisé. L'estimateur mesure une discontinuité datée à la pose : un effet qui s'installe progressivement se reporte en partie sur la pente et n'apparaît pas dans la rupture estimée (annexe 13.3). Enfin, 2,8 % des radars témoins (46 sur 1 646) ont reçu un radar neuf à moins de 500 m pendant la période (7,6 % à moins de 1 km), ce qui tire aussi l'estimation vers zéro, de moins d'un point d'après le contrôle qui exclut les zones concernées (annexe 13.6, section C). Un seul canal chiffré porte sur la précision et non sur le niveau de l'estimation : le regroupement spatial des erreurs-types, qui laisse le point inchangé (+3,4 %) mais élargit les seuils de détection (13,7 % à 15,7 % à la baisse sur la fréquence à 500 m). Ces canaux sont réels, mais ils ne peuvent pas être présentés comme expliquant le résultat : sur la fréquence, l'estimation n'est pas bornée (le seuil de 14 % ne vaut que pour une spécification que le placebo met en défaut), de sorte que l'on ne sait pas de quel zéro il faudrait rendre compte ; sur la gravité et la mortalité, la limite est la puissance. Ce que l'étude établit, c'est que le bénéfice localisé rapporté par une partie de la littérature n'est pas visible dans ces données, avec ce protocole, sur les radars français installés depuis 2019, et que l'effet global du programme n'est pas ce qu'elle mesure.

10. Conclusion

Dans ce jeu de données (radars fixes installés depuis 2019, France métropolitaine, 2017-2024, fenêtre 2018-2024 pour la gravité, proximité 500 m à 1 km), l'étude **ne détecte aucun effet localisé des radars sur les blessés hospitalisés (proxy imparfait des blessés graves, §5), sur leur total avec les tués (KSI) et sur la mortalité**. Cette absence de détection admet deux lectures. La première est qu'il n'existe pas d'effet. La seconde tient à la puissance statistique : un effet réel resterait invisible en deçà d'environ 18 à 25 % à la baisse (ou bien 22 à 34 % à la hausse) sur la gravité, et de 33 à 48 % à la baisse (50 à 94 % à la hausse) sur la mortalité, les seuils n'étant pas symétriques puisque l'échelle du modèle est logarithmique. L'hypothèse d'une baisse de mortalité, suggérée par l'Empirical Bayes mais absente du test de cassure, n'est corroborée par le test de cassure ni sur les blessés graves ni sur les KSI, six à sept fois mieux alimentés ; sur l'Empirical Bayes lui-même, une seule des quatre estimations de gravité exclut zéro (KSI à 500 m, -8,8 %, $p = 0,036$; §8, annexe 13.4) ; la calibration de l'Empirical Bayes retenue ici (annexes 13.3 et 13.4) ne corrige au demeurant pas la part de retour à la moyenne commune aux sites traités.

Sur la **fréquence des accidents**, notre étude ne permet pas de trouver un effet localisé, et ne peut pas davantage établir qu'il est nul. L'estimation dépend de la forme donnée au reflux d'accidentalité antérieur à la pose, et un test placebo ne valide aucune des formes testées : celle qui porte l'estimation publiée s'écarte de zéro de 19,2 points en médiane, toujours dans le même sens, et les seuils de détection du test lui-même, de l'ordre de 25 % pour cette spécification, ne permettent d'en certifier aucune autre (annexes 13.6 et 13.7). Le signe de l'effet local sur la fréquence n'est pas déterminé par ces données.

Une précision de cadrage s'impose sur ce qui est mesuré. Les zones témoins étant elles-mêmes des abords de radars, mis en service avant 2019, la grandeur estimée est l'effet localisé **marginal** d'un radar supplémentaire dans un réseau déjà densément équipé ; et pour deux tiers du groupe traité, l'événement daté est le remplacement d'un radar présent sur le site installé avant 2017, le plus souvent entre 2005 et 2012, non une première installation (§5), ce qui rapproche la grandeur estimée de l'effet du renouvellement d'un équipement autant que de celui de son arrivée. Ce n'est ni l'effet du programme de contrôle automatisé, ni la dissuasion générale qu'il peut exercer sur l'ensemble du réseau : celle-ci s'exerce sur les deux groupes et sort de la comparaison par différence. Ces résultats ne disent donc rien de ce qu'il adviendrait en l'absence de contrôle automatisé.

Ce que l'étude établit en revanche sans dépendre d'aucune hypothèse sur la forme de la tendance, c'est que **la baisse de 30 % observée en comparaison brute avant/après ne mesure pas l'effet des radars**. Le placebo le démontre directement : dans une fenêtre antérieure à toute pose réelle, une comparaison sans correction de tendance produit déjà un écart significatif dans six configurations sur sept, parce que l'accidentalité des sites des cohortes 2022-2024, seules à disposer d'une antériorité suffisante pour ce test, monte jusqu'à l'année qui précède la pose. C'est la contribution principale de ce travail.

11. Limites

Ces limites se rangent en quatre familles : ce que les données mesurent et avec quelle précision (11.1), la comparabilité des groupes et l'identification de l'effet (11.2), la puissance statistique et l'inférence (11.3), les interventions concomitantes et le contexte de la période (11.4). Chacune est énoncée avec le contrôle qui la lèverait lorsqu'il est réalisable sur les données existantes.

11.1 Ce que les données mesurent, et avec quelle précision

- **Profondeur temporelle.** La géolocalisation fiable ne remonte pas avant 2017, année de généralisation des tablettes chez les forces de l'ordre. Cela limite la longueur des périodes « avant » et empêche d'étudier la première vague de radars et de données ONISR (2005-2010). Pour la gravité (fenêtre 2018-2024), les radars installés dès 2019 n'offrent qu'une seule année d'observation avant traitement.
- **Précision de la position des accidents, distincte de la couverture du géocodage.** L'offset $\log(c_t)$ corrige la part d'accidents géolocalisés chaque année (§7.3), c'est-à-dire la complétude du géocodage, et non son exactitude. Aucune mesure de l'erreur de position des points BAAC n'est publiée, et les fichiers ne permettent pas de l'estimer : l'hypothèse d'une localisation exacte à l'échelle du disque de 500 m est donc invérifiable avec ces données. Sa portée est directe. Si l'erreur est indépendante du traitement, elle mélange les zones et atténue vers zéro tout effet localisé, ce qui rend les seuils de détection publiés optimistes dans une proportion que ces données ne bornent pas : c'est une explication concurrente d'un effet localisé non identifié. Si elle varie dans le temps (généralisation de la saisie sur tablette à partir de 2017, changement de format des coordonnées selon les millésimes) ou entre milieux (rase campagne côté traité, carrefours urbains côté témoin), elle n'est pas absorbée par les effets fixes d'année. Deux diagnostics descriptifs la documenteraient sans donnée nouvelle : la distribution annuelle des distances accident-radar à l'intérieur du disque de 1 km, et la part annuelle d'accidents à coordonnées suspectes (valeurs nulles, arrondies ou dupliquées). Ils n'ont pas été conduits ici, et le premier resterait un indice et non un test, un profil décroissant pouvant aussi refléter la géométrie du réseau.
- **Complétude du géocodage entre milieux.** La correction de couverture (87 % en 2017, 97 % en 2018, quasi 100 % ensuite, §7.3) est un facteur national appliqué identiquement aux deux groupes, alors que ceux-ci diffèrent par leur milieu (33,4 % des accidents hors agglomération côté traité, 17,6 % côté témoin, annexe 13.6, section D) : si la généralisation

du géocodage avait été plus lente hors agglomération, les zones traitées auraient été relativement sous-comptées en début de période, ce qui aurait pu contribuer au profil pré-installation croissant. Cette hypothèse a été mesurée (étape 14 du programme de reproduction) : la part d'accidents à coordonnées exploitables est presque identique entre milieux, et légèrement plus élevée hors agglomération (89,2 % contre 88,1 % en 2017, 97,0 % contre 96,7 % en 2018, 100 % ensuite). Pondérée par la composition des deux groupes, cette différence produit un écart de niveau de +0,20 % en 2017 et de +0,05 % en 2018, soit 0,5 % de l'amplitude du profil pré-installation entre -7 et -5 ans (0,384 point sur l'échelle logarithmique) ; la même pondération par catégorie de route donne -0,80 % en 2017 et +0,24 % en 2018 : le signe dépend de la dimension retenue, mais l'ordre de grandeur ne varie pas, inférieur à 1 % sur deux années ; à titre de comparaison, l'écart entre la catégorie de route la moins bien couverte et la mieux couverte en 2017 est de 73,0 % (autoroutes) contre 91,7 % (départementales). La couverture zone par zone reste, elle, invérifiable : un accident non géocodé n'est attribuable à aucun radar, et la stratification par milieu et par catégorie de route en est un substitut, non la mesure.

- **Indicateur blessés graves non labellisé.** La saisie de la gravité d'un blessé a changé en 2018, et l'indicateur n'est plus labellisé depuis 2019 (§5) ; l'analyse de gravité est donc restreinte à la fenêtre de saisie homogène 2018-2024, à interpréter avec cette réserve.
- **Vie du parc et statut opérationnel.** Deux limites tiennent au fichier des radars. D'une part, ce fichier est une photographie du parc à sa date de publication : un radar installé après 2019 puis déposé ou déplacé avant cette date n'y figure plus, ce qui peut introduire un biais de survie dans le groupe traité ; l'ampleur de ce biais est vérifiable en croisant les 1 521 installations retenues avec les comptages annuels des documents budgétaires CCSR. D'autre part, l'étude mesure l'effet de l'**installation** d'une cabine, non de son **fonctionnement effectif** : un radar vandalisé (part des radars détruits, dégradés, attaqués ou neutralisés évaluée à environ 75 % en mars 2019, §2.6) ou une tourelle leurre (§2.3) reste compté comme installé. L'effet estimé s'interprète donc en « intention de traiter » (présence du dispositif, actif ou non), qui est la grandeur pertinente pour la décision publique, mais qui minore mécaniquement l'effet propre d'un radar effectivement opérationnel. L'ampleur de cette dilution n'est pas quantifiable ici : la part des tourelles effectivement équipées d'un dispositif de mesure n'est pas publiée, et le fichier des radars ne porte ni date de dépose ni état de fonctionnement. Les seuls ordres de grandeur disponibles sont ceux du parc entier, dont la part en fonctionnement est passée d'environ 88 % à 81 % en 2024 (§2.6 et sa note), et le pic de dégradation de la fin 2018 (§2.6). L'estimation publiée porte donc sur la présence du dispositif, et non sur son fonctionnement : elle ne se transpose pas telle quelle à un parc intégralement opérationnel. Pour un estimateur de saut, une mise en service effective postérieure à

l'installation ne se réduit d'ailleurs pas à une dilution de l'effet : elle déplace la date de la discontinuité recherchée, ce qui pénalise davantage ce type d'estimateur qu'un estimateur de moyenne postérieure au traitement (annexe 13.3). Cette réserve porte sur le canal de la sanction (avis de contravention, apprentissage individuel, information indirecte) et non sur la dissuasion visuelle immédiate : une tourelle non équipée étant indiscernable d'une tourelle équipée pour le conducteur qui passe, l'effet local sur les vitesses pratiquées, et donc sur l'accidentalité mesurée ici, ne devrait guère en dépendre.

- **Absence de mesure d'exposition au trafic au niveau du site.** Les comptages ne sont normalisés par aucun volume de circulation local : l'offset des modèles est la couverture du géocodage (§7.3), et le trafic n'intervient qu'en tendance nationale, dans une variante de robustesse (annexe 13.4). Aucune source ne fournit un trafic annuel comparable pour les abords des 3 167 radars étudiés. Trois conséquences. Les zones traitées comptent environ trois fois moins d'accidents que les zones témoins, écart de niveau absorbé par l'indicatrice de groupe mais dont la part imputable au trafic est inconnue. Une évolution locale de circulation (déviation, chantier, mise en service d'un contournement, reprise différenciée après 2020) n'est absorbée ni par les effets fixes d'année, qui sont nationaux, ni par la pente propre aux traités, qui est linéaire. Enfin, la fonction de performance de sécurité employée dans l'Empirical Bayes est calée sur la période antérieure des sites traités et non sur une mesure de trafic, contrairement à la pratique canonique (§3.3.3, annexe 13.3) : ce n'est pas une fonction de performance calibrée sur l'exposition, et elle ne corrige pas ce que corrigerait une exposition mesurée. Le *Bilan de la circulation* du SDES distingue les autoroutes, les routes nationales et les autres routes, mais ne comporte pas de ventilation régionale ou départementale : une tendance différenciée par catégorie de réseau serait calculable, elle n'a pas été estimée ici, et une exposition au niveau du site ne l'est pas.

11.2 Comparabilité des groupes et identification de l'effet

- **Comparabilité des groupes.** Zones à radars récents et anciens diffèrent en niveau de base : les zones traitées comptent environ trois fois moins d'accidents que les zones témoins (coefficient de niveau $\exp(-1,13) \approx 0,32$, annexe 13.1). Cet écart de niveau est absorbé par l'indicatrice de groupe ; le plan repose donc sur une hypothèse de tendances parallèles, par nature non testable directement. Le profil pré-installation estimé à composition constante sur les cinq années qui précèdent la pose, sur les cohortes 2022-2024 (414 zones traitées sur 1 002, seules à disposer de cinq à sept années antérieures ; annexes 13.6 et 13.7) montre que les tendances ne sont **pas** parallèles avant la pose : l'accidentalité des zones traitées monte jusqu'à l'année précédant l'installation. C'est ce qui motive l'ajout d'une pente spécifique aux traités, l'effet n'étant identifié que par la

rupture, et c'est aussi ce qui rend cette pente mal spécifiée, aucune tendance monotone ne pouvant épouser une trajectoire qui monte puis redescend. Cette pente porte en outre l'essentiel de l'écart entre la comparaison sans pente (-28,9 %) et l'estimation publiée (+3,4 %) : cet écart est le produit de la pente par le décalage moyen du temps d'événement, à un point près (étape 17 du programme de reproduction). Or, en l'absence d'effet fixe de zone et avec des indicatrices d'année saturées, elle n'est pas identifiée par la trajectoire propre d'un site : elle capte pour partie le gradient de niveau entre cohortes de pose, et ces données ne permettent pas d'en séparer les deux composantes (annexe 13.6, section E). Cet écart de niveau soulève par ailleurs une difficulté. Si les radars étaient posés de façon strictement réactive sur des points noirs, comme l'indique le §3.1, les sites récents devraient ressembler aux anciens et non être trois fois moins accidentogènes. Deux lectures sont possibles, aux conséquences opposées : soit la doctrine d'implantation 2019-2024 n'est plus principalement réactive (remplacement de cabines vandalisées, densification du parc, zones de contrôle), auquel cas le retour à la moyenne est plus faible qu'on ne le suppose ; soit les sites récents sont d'une autre nature que ceux des vagues antérieures, ce qui contribuerait à expliquer l'écart avec les évaluations françaises portant sur la première vague (§3.6). Un élément factuel étaye la première lecture pour la plus grande part du groupe : 68 % des radars datés 2019-2024 remplacent un radar présent au même endroit en octobre 2018 (§5), et pour ces sites le choix du lieu remonte le plus souvent à 2005-2012, non à un pic récent. Un élément va en sens inverse pour le reste : le test placebo (annexe 13.7) établit que l'accidentalité des cohortes 2022-2024, dont la moitié environ sont des sites nouveaux, **monte** significativement jusqu'à l'année qui précède la pose, ce qui atteste une composante de sélection réactive réelle, indépendamment du niveau de base. L'hypothèse résiduelle porte sur la forme, supposée linéaire, de ce reflux : examinée sur la fréquence, elle est difficile à contrôler sur les événements rares.

- **Nature du traitement : renouvellement plutôt qu'installation.** Le fichier des radars datant l'équipement et non le site, l'événement daté est, pour 1 030 des 1 521 radars du groupe traité, le remplacement d'un radar présent le plus souvent depuis 2005-2012, avec le cas échéant une interruption de service que rien ne permet de mesurer (§5). Pour ces sites, la période « avant » est déjà sous radar et aucun effet d'arrivée n'est attendu, ce qui tire l'estimation d'ensemble vers zéro (§9) ; l'effet d'un renouvellement (tourelle plus visible, contrôle dans les deux sens, remise en service après dégradation) n'est pas nul a priori mais n'est pas celui que la littérature évalue. Une estimation séparée sur les seuls sites nouveaux (472 radars, 327 zones à 500 m) est possible et a été calculée à titre de diagnostic (annexe 13.6, section E) ; elle n'est pas publiée comme résultat, parce que la spécification y échoue au placebo plus nettement que sur l'ensemble et que le contraste

entre les deux sous-groupes change de sens selon l'estimateur (la spécification publiée, qui comporte une pente, les sépare dans un sens ; le même modèle sans pente les sépare dans l'autre ; une différence de différences à effets fixes de cohorte ne les sépare plus), alors que l'Empirical Bayes et l'estimateur de Callaway-Sant'Anna ne distinguent pas les deux sous-groupes.

- **Confusion entre statut du site et millésime de pose.** Les remplacements se concentrent en 2019-2020 (88 % et 91 % des millésimes) et les sites nouveaux en 2021-2024 : tout contraste entre les deux sous-groupes mêle la nature du site à l'année de pose, donc à la durée de suivi (1,9 année après la pose en moyenne pour les nouveaux, 3,2 pour les rééquipés) et au niveau d'accidentalité des cohortes. Ce partage n'était pas prévu au protocole et rien ne permet, avec ces données, d'attribuer au seul statut du site un écart entre sous-groupes.
- **Stabilité supposée de l'effet propre des radars témoins.** Les zones témoins sont des abords de radars mis en service avant 2019 : l'estimation suppose que leur effet propre est stabilisé sur 2017-2024. Si cet effet s'érode, l'accidentalité des zones témoins dérive à la hausse, les effets fixes d'année, identifiés principalement par les témoins, transportent cette dérive dans le contrefactuel des zones traitées, et la rupture estimée est déplacée vers la baisse ; une consolidation de cet effet la déplacerait vers la hausse. La littérature française citée au §3.6 documente précisément un effet décroissant dans le temps (Roux & Zamora, 2013 ; Carnis & Blais, 2013), sur des radars d'une autre génération. Le fichier des radars portant la date de mise en service, nous avons restreint les témoins aux radars mis en service avant 2013 (923 des 1 274 zones témoins), puis avant 2010 (343 zones) : l'estimation passe de +3,4 % à +3,9 % puis +3,8 % à 500 m, de +4,9 % à +5,0 % puis +5,2 % à 1 km, les seuils de détection restant inchangés au dixième de point près (annexe 13.6, section C). Un reflux résiduel des témoins les plus récents ne porte donc pas le résultat ; ce contrôle ne dit rien de la spécification elle-même.
- **Composition technologique des groupes, et absence de stratification par type de radar.** Tous les radars fixes sont traités uniformément, alors que la composition des deux groupes est en pratique disjointe (annexe 13.6, section D) : le groupe traité est constitué à 87 % de radars « tourelles » (1 328 sur 1 521) et ne contient quasiment aucune cabine fixe classique, tandis que le groupe témoin comprend 637 cabines fixes, 529 radars de feu rouge, 339 radars discriminants, 71 passages à niveau et 70 radars tronçon, ces deux derniers types ayant une zone d'influence qui se prête mal à un disque de 500 m. L'étude porte donc de fait, pour l'essentiel, sur les tourelles mises en service depuis 2019, et l'hypothèse de tendances parallèles est formulée entre deux populations qui diffèrent par la technologie autant que par l'ancienneté. Un contrôle restreignant les traités aux tourelles et les témoins aux cabines fixes et discriminants donne +0,6 % ($p = 0,91$) : cette

différence de composition ne déplace pas l'estimation ponctuelle au-delà de ce que ce contrôle peut distinguer, sa sensibilité étant de 14,8 % à la baisse et de 17,3 % à la hausse (11 429 lignes, intervalle de confiance [-10,0 % ; +12,6 %]). Les radars de feu rouge, les radars tronçon (contrôle de vitesse moyenne sur plusieurs kilomètres), les radars « tourelle » ou les radars autonomes ont des mécanismes et des zones d'influence différents, non distingués dans l'estimation principale. Les premiers ont un effet propre documenté sur l'accidentalité, à la baisse pour les collisions à angle droit et possiblement à la hausse pour les collisions par l'arrière (Cohn et al., 2020 ; Gallagher & Fisher, 2020, qui, après l'arrêt par référendum du programme de Houston, observent une redistribution des types de collisions sans variation du total) : leur présence dans le groupe témoin est prise en compte par le contrôle en groupes homogènes rapporté ci-dessus. De même, aucune stratification par vitesse limite (VMA), par milieu (agglomération ou rase campagne) ou par catégorie de route n'est proposée : ces sous-groupes sont possibles avec les données disponibles (la VMA figure au fichier des radars), mais chaque découpage divise la puissance ; les seuils détectables, d'environ 14 % sur la fréquence pour l'ensemble, s'élargiraient sensiblement à chaque découpage (déjà 18,6 à 18,9 % à la baisse à 500 m pour les deux seuls sous-groupes examinés à l'annexe 13.6, section E), et davantage sur la gravité. Une telle analyse exigerait des sous-groupes définis a priori et un contrôle de la multiplicité des tests. Une comparaison purement descriptive des deux groupes sur ces trois dimensions, qui ne consomme aucune puissance statistique, est en revanche publiée à l'annexe 13.6, section D.

- **Univers de l'estimation défini sur la variable expliquée.** Les zones retenues sont celles qui comptent au moins un accident géolocalisé sur la fenêtre (annexe 13.4), filtre qui porte sur la grandeur estimée elle-même et s'applique de façon asymétrique : 1 002 des 1 521 zones traitées et 1 274 des 1 646 zones témoins. Le contrôle qui le lève, l'univers complet comptages nuls inclus, donne +6,6 % ($p = 0,23$) au lieu de +3,4 %, soit 3,2 points de déplacement, les seuils de détection restant inchangés (annexe 13.6, section C).
- **Témoins partiellement exposés.** Un radar neuf a été installé pendant la période à moins de 500 m de 2,8 % des radars anciens (46 sur 1 646) et à moins de 1 km de 7,6 % d'entre eux (125) : ces zones témoins ont donc été marginalement « traitées » à leur tour, ce qui tirerait l'estimation vers zéro. Leur exclusion déplace l'estimation de +3,4 % à +3,2 % (témoins exposés à moins de 500 m) puis +2,8 % (à moins de 1 km) au rayon de 500 m, et de +4,9 % à +4,7 % au rayon de 1 km, seuils inchangés (annexe 13.6, section C) : cette contamination, telle que le fichier des radars permet de la borner, ne porte pas le résultat.
- **Report spatial.** Si un radar déplace une partie des accidents au-delà de la zone étudiée (freinages tardifs, ré-accélérations en aval), l'estimation locale en serait biaisée vers la baisse ; un déplacement en sens inverse, depuis l'extérieur vers l'intérieur du disque, la

biaiserait vers la hausse. Un déplacement interne au disque, lui, ne modifie pas le comptage de la zone et ne déplace pas l'estimation. La méta-analyse de Høye (2014), qui traite explicitement les sauts de kangourou (freinage puis ré-accélération), observe ce comportement de conduite sans détecter d'effet défavorable sur les accidents, et juge la migration d'accidents possible mais non courante. Le rayon de contrôle à 1 km capte une partie de ces reports, sans exclure un déplacement au-delà. Une estimation portant sur le seul anneau compris entre 500 m et 1 km, comparée à celle du disque, documenterait directement la question : un écart de signe entre les deux signifierait un déplacement, leur concordance l'écarterait. Ce contrôle est réalisable sur les données existantes ; il n'a pas été conduit ici.

- **Zones de contrôle signalées globalement (25 km).** Sur certains axes, la signalisation n'est plus localisée à chaque cabine : l'hypothèse d'une influence concentrée à 500 m-1 km y est plus faible. Malheureusement, les données de l'ONISR ne permettent pas d'identifier ces cas.

11.3 Puissance statistique et inférence

- **Événements rares.** La mortalité est un événement statistiquement rare : chaque site n'enregistre qu'un très petit nombre de tués, ce qui rend la puissance statistique intrinsèquement faible pour cet indicateur.
- **Corrélation spatiale entre zones voisines.** Chaque accident est rattaché à son radar le plus proche (aucun double compte ; lorsque deux zones se recouvrent, la zone effective est le disque tronqué à mi-distance du radar voisin). La surface effectivement observée dépend donc de la densité locale du parc, sans que cette dépendance ait été quantifiée ici ; elle est constante sur la période, et l'indicatrice de groupe n'en absorbe que la moyenne par groupe, le modèle ne comportant pas d'effet fixe de zone. Une seconde limite tient au voisinage lui-même : deux radars distants de quelques centaines de mètres partagent la même route et les mêmes aléas de trafic : leurs comptages ne sont pas indépendants. Les erreurs-types, regroupées par zone, captent la corrélation dans le temps au sein d'une même zone, pas cette corrélation entre zones voisines ; elles peuvent donc être légèrement sous-estimées. Un regroupement approché par maille spatiale ($\approx 1^\circ$ de latitude et de longitude) les augmente d'environ 15 % et porte le seuil détectable sur la fréquence à 500 m de 13,7 % à 15,7 % à la baisse (de 15,9 % à 18,7 % à la hausse ; annexe 13.6, section C) : les seuils publiés au §8 et au §9 sont donc de légers minorants. Un regroupement par département constituerait un contrôle complémentaire.

11.4 Interventions concomitantes et contexte de la période

- **Co-intervention majeure : l'abaissement à 80 km/h et son retour partiel.** Le passage généralisé à 80 km/h sur les routes bidirectionnelles sans séparateur est intervenu le 1^{er} juillet 2018 ; à partir de 2019, une soixantaine de départements sont revenus à 90 km/h, à des dates échelonnées. Cette co-intervention est de grande ampleur et différenciée dans l'espace et dans le temps : elle porte précisément sur les catégories de routes où sont implantés les radars fixes et chevauche à la fois la période de référence (2017-2018) et la période de traitement (2019-2024). Les effets fixes d'année du modèle absorbent la composante nationale de cette mesure, mais non sa composante départementale : si le retour à 90 km/h est corrélé à des caractéristiques départementales elles-mêmes corrélées à l'implantation des radars récents, un biais résiduel subsiste. Un contrôle par effets fixes département $\times$ année, ou par exclusion des départements repassés à 90 km/h, lèverait cette réserve ; il n'a pas été conduit ici.
- **Choc de circulation de 2020-2021.** La fenêtre d'observation contient l'effondrement puis le rebond du trafic liés à la crise sanitaire. Ce choc n'a pas été proportionnel selon le réseau : d'après le *Bilan de la circulation* du SDES, la circulation de 2020 recule de 22,1 % sur les routes nationales, de 19,3 % sur les autoroutes et de 16,9 % sur les autres routes par rapport à 2019, et le rebond de 2021 reste inégal (3,8 % sous le niveau de 2019 sur les autoroutes, 11,9 % en dessous sur les autres routes). Les effets fixes d'année, identifiés principalement par les témoins dont 32 % sont des radars de feu rouge donc urbains, n'absorbent ce choc qu'à la condition qu'il agisse de façon proportionnellement identique sur les deux groupes, ce que ces écarts par réseau ne garantissent pas. Deux difficultés s'y ajoutent : 2020 et 2021 appartiennent à la période postérieure à la pose pour les cohortes 2019 et 2020, et à la période antérieure pour les cohortes 2022 à 2024, de sorte que le choc n'entre pas au même titre dans les deux branches du modèle ; et l'étude ne dispose d'aucune mesure d'exposition au trafic au niveau du site pour le neutraliser (point suivant). L'annexe 13.6, section A, en donne la sensibilité : l'exclusion de la seule année 2020 déplace l'estimation de +3,4 % ($p = 0,52$) à +13,4 % ($p = 0,044$). Dans le test placebo, les deux configurations hors 2020 produisent les écarts à zéro les plus grands de la spécification retenue (-30,8 % et -44,4 %, annexe 13.7).
- **Co-interventions locales non documentées.** L'installation d'un radar s'accompagne parfois d'autres aménagements (réfection de chaussée, marquage, éclairage, reprise d'un carrefour) qu'aucune donnée disponible ne recense. L'effet mesuré est donc celui de l'ensemble « radar + aménagements éventuels ». Ces co-interventions étant censées réduire l'accidentalité, leur présence éventuelle orienterait l'estimation vers la baisse : elle

ne peut donc pas expliquer l'absence d'effet détectable sur la gravité et sur la mortalité. Sur la fréquence, où la spécification elle-même n'est pas validée (§8, annexe 13.7), aucune borne de ce type ne peut être posée.

- **Progrès concomitants de sécurité des véhicules et des infrastructures.** La période 2017-2024 a vu se poursuivre la diffusion des équipements de sécurité dans le parc roulant et la mise en œuvre de programmes d'infrastructure. Le contrôle électronique de stabilité est obligatoire sur tous les véhicules neufs immatriculés depuis le 1er novembre 2014, et l'appel d'urgence embarqué (eCall) sur les nouveaux types de véhicules légers depuis le 31 mars 2018 ; leur diffusion dans le parc roulant suit le rythme de renouvellement de celui-ci, de l'ordre de 4,5 % par an (1,76 million d'immatriculations neuves en 2024 pour environ 39 millions de voitures en circulation au 1er janvier 2024, d'un âge moyen de l'ordre de 11 ans ; SDES, 2024 et 2025). Il faut distinguer deux cas, de portée très différente pour l'identification. **Premier cas :** ces progrès portent sur le parc automobile national ou sur des programmes d'infrastructure sans lien avec l'implantation des radars, et sont donc **communs aux zones traitées et aux zones témoins**. Le modèle de l'annexe 13.3 comporte des effets fixes d'année γ_y , identifiés principalement par les témoins, qui absorbent par construction toute évolution nationale de cette nature ; la pente propre aux traités k_{zt} absorbe en outre la composante linéaire d'une évolution qui serait ressentie différemment par les deux groupes. Cette absorption n'est cependant exacte que si l'évolution agit de façon **proportionnellement identique** sur les deux groupes, hypothèse qui n'est pas garantie : les groupes diffèrent par leur composition (87 % de tourelles côté traité, 529 radars de feu rouge côté témoin, voir ci-dessus), donc par leur milieu et leur type de route, et un gain de sécurité véhicule concentré sur les sorties de route ne se répartit pas de la même façon en rase campagne et en carrefour urbain. La composante non linéaire d'un tel écart n'est pas absorbée. **Second cas :** une co-intervention **différenciée**, concentrée sur les abords des radars récents et datée à la pose, du type réfection de chaussée ou reprise de carrefour accompagnant l'installation d'une cabine. Celle-là est indiscernable de θ , puisqu'elle est discrète à la même date. Elle est signalée ci-dessus sous « co-interventions locales non documentées », avec sa conséquence : orientant l'estimation vers la baisse, elle ne peut pas expliquer l'absence d'effet détectable sur la gravité et sur la mortalité, et ne fonde aucune borne sur la fréquence. Son vecteur plausible est la reprise d'infrastructure : réfection d'une couche de roulement, traitement d'un obstacle latéral ou reprise d'un carrefour programmés à l'occasion de l'implantation d'une cabine. Aucune source ne quantifie la fréquence de cette concomitance, qui demeure donc une limite non mesurable et non un contrôle. Enfin, le règlement (UE) 2019/2144, souvent invoqué à propos de l'adaptation intelligente de la vitesse, s'applique à tous les véhicules neufs depuis le 7 juillet 2024 seulement :

compte tenu du rythme de renouvellement du parc, de l'ordre de 4,5 % par an, il ne concerne sur la fenêtre d'observation qu'une fraction négligeable du parc roulant et ne constitue pas une co-intervention de la période.

Précision de portée. Les progrès de sécurité des véhicules et des infrastructures expliquent une part de la baisse tendancielle nationale de l'accidentalité. C'est précisément pourquoi les évaluations agrégées qui attribuent cette baisse au contrôle automatisé sont exposées au biais de variable omise, point développé aux §2.5 et §3.6. Réciproquement, la présente étude ne décompose pas la tendance nationale, ne mesure pas la contribution respective du parc automobile, des infrastructures et du contrôle automatisé à cette baisse, et ne permet donc en aucune façon de conclure que la baisse de la mortalité viendrait des véhicules plutôt que des radars. Elle mesure un effet **localisé** à l'abord des radars installés depuis 2019, relativement à d'autres zones à radars, sur la période 2017-2024.

Financement, conflits d'intérêts et disponibilité des données

Conflits d'intérêts. L'auteur, Jean-Philippe Casoni, et l'association Réalités Routières déclarent n'avoir aucun conflit d'intérêts.

Origine du travail. L'étude procède d'une initiative propre de l'association. Elle n'a été ni commandée, ni suggérée, ni suscitée par un tiers, quel qu'il soit. Le choix du sujet, du protocole et du calendrier de publication relève de la seule association.

Financement. L'étude est financée exclusivement par des dons privés. Elle n'a reçu aucun financement public, ni d'un parti politique, ni d'une autre association. Aucune entreprise du secteur du transport routier ou d'un secteur assimilé n'a contribué à son financement. Aucun donateur n'est intervenu dans la conception du protocole, l'analyse des données, la rédaction ni la décision de publier.

Disponibilité des données et du code. L'étude repose exclusivement sur des données publiques : fichiers BAAC de l'ONISR, fichier des radars automatiques de data.gouv.fr (état actuel et extraction d'octobre 2018), *Bilan de la circulation* du SDES. Plusieurs documents permettent de reproduire nos résultats :

- `reproduction_etude_radars.ipynb` : Il reproduit les chiffres des deux résumés, du §2.6, du §3.7, du §5, du §6, du §7.2, du §7.3, du §8, du §9, du §10, du §11 et des annexes 13.1 à 13.7.
- `pipeline_reproduction.md` en donne une transcription commentée intégrale : les dix-sept étapes, code repris à l'identique du programme et valeurs attendues reportées sous chaque bloc ; il se termine par une table de traçabilité qui associe chaque chiffre publié à l'étape qui le produit. Les grandeurs qui ne sont pas des sorties directes y sont identifiées comme telles : celles qui se déduisent d'une sortie imprimée par une opération arithmétique portent la mention « provient de », qui nomme cette sortie et l'étape concernée, et celles qui proviennent d'une vérification conduite hors programme portent la mention « aucune ».

Le programme est publié à l'adresse realitesroutieres.fr/etude-radars-efficacite-2026 et les jeux de data.gouv.fr sont mis à jour en continu et les millésimes BAAC font l'objet de révisions après publication, de sorte qu'une extraction téléchargée ultérieurement ne redonne pas nécessairement les effectifs de contrôle imprimés en tête du programme. La liste des fichiers sources attendus figure en tête du programme ; leurs références sont données au §12.3.

À propos de Réalités Routières

Réalités Routières · Observatoire des mobilités est une association d'intérêt général ayant pour objectif d'apporter un regard neuf et indépendant sur les politiques publiques de mobilité. Site : realitesroutieres.fr ; l'ensemble des sources et scripts de cette étude est disponible sur realitesroutieres.fr/etude-radars-efficacite-2026.

12. Sources et références

12.1 Notes de bas de page (§2, contexte historique ; §4, effets négatifs et externalités)

- 1 Inauguration du premier radar automatique fixe sur la RN20 à La Ville-du-Bois (Essonne) le 27 octobre 2003 par Nicolas Sarkozy (ministre de l'Intérieur) et Gilles de Robien (ministre de l'Équipement et des Transports), en présence de Rémy Heitz (délégué interministériel à la sécurité routière) ; premiers avis de contravention envoyés à compter du 1er novembre 2003. ONISR, *La sécurité routière en France* (bilans annuels). Rétrospectives de presse : *L'argus*, « L'histoire secrète des radars automatiques, à l'occasion des 20 ans de la première cabine », 27 octobre 2023 ; France 3 / ICI, « Il y a 20 ans, l'inauguration du premier radar automatique par le ministre picard Gilles de Robien », octobre 2023.
- 2 ONISR, *La sécurité routière en France, bilan de l'accidentalité de l'année* (éditions annuelles), Observatoire national interministériel de la sécurité routière : séries de mortalité routière 2001-2004.
- 3 Parc fin 2008 : rapport annuel de performances 2008 de la mission *Contrôle et sanction automatisés des infractions au code de la route*, annexe au projet de loi de règlement des comptes pour 2008, programme 751 « Radars », tableau « Données physiques », p. 35 (1 473 dispositifs fixes et 827 mobiles installés fin 2008, soit 2 300 ; le marché des radars de franchissement de feu rouge n'étant notifié qu'au dernier trimestre 2008, ces 1 473 sont tous des radars de contrôle de vitesse). Évolution 2004-2019 : Cour des comptes, *La politique publique de sécurité routière*, juin 2021, p. 37 (257 dispositifs en 2004, 2 300 en 2008, 4 097 en 2013, 4 428 en 2018, tous types confondus) et tableau n° 1, p. 38 (par type, d'après les rapports annuels de performance : radars fixes 1 473 en 2008, 2 203 en 2013, 2 051 en 2016, 1 989 en 2018, 1 291 en 2019 ; total 4 094 en 2019, dont 375 tourelles). Le parc contemporain et sa cible sont détaillés par type d'équipement dans le projet annuel de performances pour 2025 (« bleu » budgétaire, p. 28 à 31, détail par type au §12.3), d'où provient la cible de 4 160 équipements donnée au §2.2. Les effectifs de radars fixes de contrôle de vitesse et ceux du parc tous dispositifs sont deux séries distinctes, précisées comme telles à chaque mention. Comparaison internationale pour l'Italie : recensement officiel du ministère italien des Infrastructures et des Transports (MIT), liste nationale des *autovelox* publiée le 28 novembre 2025 (~ 11 000 dispositifs informellement recensés, dont ~3 150 à 3 625 homologués et actifs).
- 4 *Utilisation par l'Agence de financement des infrastructures de transport de France (AFIT France) et par les collectivités territoriales des recettes issues du compte d'affectation spéciale « Contrôle de la circulation et du stationnement routiers »*, annexe générale au projet de loi de finances (« jaune » budgétaire), Direction du budget, éditions annuelles ; base légale : article 160 de la loi de finances pour 2017. Recettes des amendes forfaitaires et majorées, 2023 et 2024 ; cumul depuis 2003. Ce document est distinct du projet annuel de performances du CAS CCSR (« bleu » budgétaire, cité au §12.3) et des notes d'exécution budgétaire de la Cour des comptes : les trois sont d'auteurs et de statuts différents.
- 5 Cour des comptes, *Note d'analyse de l'exécution budgétaire 2024 : Contrôle de la circulation et du stationnement routiers*, avril 2025 (part des dépenses présentant un lien avec la sécurité routière ; ventilation de l'affectation). <https://www.ccomptes.fr/sites/default/files/2025-04/NEB-2024-Contrôle-circulation-et-stationnement-routiers.pdf>

- 6 J.-G. Padioleau, *La société du radar : l'État sarkozyste au Kärcher*, Paris, L'Harmattan, décembre 2005 ; voir aussi sa tribune dans *Libération* du 6 janvier 2004.
- 7 Ligue de défense des conducteurs, *Dossier de l'étude vitesse*, juillet 2013 (étude réalisée par le cabinet SCM), contestant le lien vitesse/mortalité.
- 8 International Transport Forum (ITF/OCDE), *Speed and Crash Risk*, Paris, ITF, 2018.
- 9 Passage généralisé aux 80 km/h (1er juillet 2018) puis mouvement des « gilets jaunes » (fin 2018) : vague de dégradations des radars. Christophe Castaner, ministre de l'Intérieur, évalue à environ 75 % la part des radars détruits, dégradés, attaqués ou neutralisés (déclaration de mars 2019, rapportée par l'AFP et *Le Parisien*). Coût, selon la réponse du ministre de l'Intérieur à la question écrite du Sénat n° 08466 (JO Sénat, 6 février 2020) : réparations estimées à ~33 M€ en 2018 et ~36 M€ en 2019, et manque à gagner d'environ 300 M€ lié à la destruction d'un millier de radars ; certaines estimations de presse portent la perte de recettes jusqu'à ~660 M€ sur deux ans (*Les Échos*). Reprise lors de la colère agricole de 2024 (radars bâchés, coupés ou brûlés) : France Bleu, « Colère des agriculteurs : des radars recouverts de bâches en Lozère », novembre 2024 ; part des radars en fonctionnement passée d'environ 88 % à 81 % et recettes 2024 en repli d'environ 76 M€ (documents budgétaires CCSR ; Cour des comptes, avril 2025).
- 10 Adhésion de l'opinion au contrôle automatisé : Ifop pour *Dimanche Ouest France*, *L'adhésion aux nouvelles mesures de sécurité routière*, résultats détaillés, décembre 2011 (étude n° 110010, échantillon de 1 002 personnes de 18 ans et plus, terrain du 6 au 8 décembre 2011), qui rappelle les enquêtes Ifop pour le même titre d'octobre 2003 (« Les premiers radars automatiques [...] Vous personnellement, approuvez-vous cette mesure ? », 86 % d'approbation) et des 14-15 octobre 2004 (« Approuvez-vous ou désapprouvez-vous la poursuite de l'installation des radars automatiques sur les routes de France ? », 72 %), contre 36 % d'approbation en décembre 2011 pour 400 radars supplémentaires non signalés par des panneaux ; le Sénat (rapport d'information n° 644, 2016-2017, avant-propos) rappelle de même que « il y a quinze ans, deux tiers des Français se prononçaient en faveur de l'instauration d'un contrôle sanction automatisé ; la tendance s'est depuis inversée ». Sentiment de « radar tirelire » : même rapport sénatorial (« selon le sondage de Harris Interactive publié le 27 juin 2014, 64 % des Français estiment que les radars automatiques servent seulement à remplir les caisses de l'État ») ; sondage Harris Interactive pour RTL et *Auto Plus*, rapporté par Europe 1, « Sondage : les radars sont des "pompes à fric" », 27 juin 2014 (64 % ; 35,5 % jugent qu'ils aident à réduire les accidents mortels). Ces chiffres mesurent une perception, non une efficacité, et l'item de 2014 est formulé de façon exclusive (« seulement »). Voir aussi le Baromètre AXA Prévention *Le comportement des Français au volant* (édition annuelle depuis 2004). La mesure française la plus complète du sentiment de « radar tirelire » est le sondage BVA du 12 au 23 mai 2014, annexé au rapport d'évaluation interministérielle de juillet 2014 (voir §12.3) : 79 % des conducteurs jugent que les actions publiques de sécurité routière ont pour objectif de « remplir les caisses de l'État », et près de 8 sur 10 que les radars sont placés là où ils rapportent plutôt que là où le danger est réel ; le même sondage relève, en sens inverse, 84 % de conducteurs estimant qu'il est encore possible de diminuer largement le nombre de tués et 71 % que ces actions réduisent réellement le nombre de victimes, et le rapport oppose à ces perceptions environ 0,8 Md€ de recettes de radars contre environ 4 Md€ de dépenses de l'État pour la sécurité routière. Sur les radars les plus verbalisateurs : palmarès annuels de presse, *L'argus*, « Le palmarès des radars qui ont le plus flashé en 2024 » (ex. A8 à Cagnes-sur-Mer, A1 au Bourget).

- 11 Fraude dite des « garages fantômes » : note de la Direction nationale de la police judiciaire (septembre 2025) évaluant à plus de 250 000 le nombre de véhicules concernés chaque année ; estimation de la Cour des comptes d'environ un million de véhicules échappant au contrôle ; question écrite à l'Assemblée nationale n° 4793. Presse : RiskAssur-hebdo, https://www.riskassur-hebdo.com/actu01/actu_auto.php?adr=2302260834 ; Auto Plus, « Garages fantômes : échapper aux PV », <https://www.autoplus.fr/actualite/garages-fantomes-echapper-aux-pv-596282.html>
- 12 Usurpation de plaque d'immatriculation (« doublette ») : procédure et charge de la contestation décrites par service-public.gouv.fr, fiche « Usurpation de plaque d'immatriculation d'un véhicule » (dépôt de plainte préalable, contestation de chaque avis auprès de l'ANTAI dans les délais impartis, justificatifs d'absence à fournir, demande d'un nouveau numéro d'immatriculation).
- 13 Le standard de la « conservation généralisée et indifférenciée » des données, développé à propos des données de connexion, sert de référence dans les débats juridiques sur la surveillance de masse : CJUE, *Digital Rights Ireland*, aff. jointes C-293/12 et C-594/12, 8 avril 2014 (invalidation de la directive 2006/24/CE sur la conservation des données) ; CJUE, *La Quadrature du Net e.a.*, aff. jointes C-511/18, C-512/18 et C-520/18, 6 octobre 2020 (une conservation généralisée et indifférenciée des données de trafic et de localisation ne peut être justifiée que par une menace grave pour la sécurité nationale). Pour les radars tronçon, l'arrêté du 13 octobre 2004 modifié, qui encadre le traitement du contrôle automatisé, impose l'effacement sous 24 heures des plaques des véhicules en règle.
- 14 Article L. 121-3 du code de la route (redevabilité pécuniaire du titulaire du certificat d'immatriculation). Conseil constitutionnel, décision n° 99-411 DC du 16 juin 1999 (en matière répressive, une présomption de culpabilité n'est admise qu'à titre exceptionnel, dès lors qu'elle n'est pas irréfragable et que les droits de la défense sont préservés) ; CEDH, *Falk c. Pays-Bas* (décision), 19 octobre 2004, req. n° 66273/01 (compatibilité de la responsabilité du titulaire de la plaque avec la présomption d'innocence, article 6 § 2) ; CEDH, *O'Halloran et Francis c. Royaume-Uni* [GC], 29 juin 2007, req. n° 15809/02 et 25624/02 (l'obligation de désigner le conducteur, sous peine d'amende, ne méconnaît pas le droit de ne pas contribuer à sa propre incrimination).
- 15 Contrôles sur place de la CNIL, septembre à décembre 2018 (procédure n° 2018-071C) : plaques complètes de véhicules en règle conservées plus de 13 mois et plaques tronquées plus de 4 ans, au lieu des 24 heures fixées par l'arrêté du 13 octobre 2004 ; messages d'infraction conservés au Centre national de traitement au-delà du plafond de 10 ans ; sécurité insuffisante (robustesse des mots de passe, traçabilité des accès, gestion des droits d'accès chez le prestataire). Mise en demeure du ministère de l'Intérieur n° MED-2019-027 du 12 novembre 2019, rendue publique le 4 décembre 2019 en raison du nombre de personnes concernées et du risque pour la vie privée ; clôture après mise en conformité, publiée le 25 mars 2021 (Légifrance, CNILTEXT000043300190). Hors ces procédures formalisées, les constats des contrôles de la CNIL ne sont pas systématiquement publiés ; la Cour des comptes audite pour sa part la gestion du dispositif (avril 2025), non le sort des données.
- 16 Le Conseil constitutionnel rappelle de façon constante qu'il ne dispose pas d'un pouvoir général d'appréciation et de décision « identique à celui du Parlement » (décision n° 74-54 DC du 15 janvier 1975, cons. 1) et ne censure qu'une disproportion manifeste. Sur l'obligation positive de protéger la vie : CEDH, *Osman c. Royaume-Uni* [GC], 28 octobre 1998, req. n° 23452/94 ; pour son application à la sécurité routière, CEDH, *Smiljanić c. Croatie*, n° 35983/14, 25 mars 2021 (violation de l'article 2, l'État n'ayant pas appliqué de manière effective son propre cadre répressif routier à un conducteur récidiviste).

- 17 Coût unitaire d'installation d'un radar automatique selon le type d'équipement : *Auto Plus*, « Combien coûte vraiment un radar ? », <https://www.autoplus.fr/actualite/combien-coute-vraiment-un-radar-1333804.html> (presse spécialisée ; les documents budgétaires du CAS CCSR ne publient pas de coût unitaire d'installation, seulement des enveloppes agrégées). Le seul barème officiel par type d'équipement identifié porte sur l'entretien et non sur l'installation : Sénat, rapport n° 644 (2016-2017), d'environ 6 600 € par appareil pour un radar fixe double sens à environ 36 000 € pour un radar autonome de chantier ; ces montants mesurent une autre grandeur et ne peuvent pas se substituer au coût d'installation.

12.2 Références scientifiques et méthodologiques

- Amancio, E. C., Gadda, T. M. C., Domingos, M. D. I., Bastos, J. T., Bonetti, G. da C., Ferreira, S. M. P., Schmitz, A. & Oviedo-Trespalacios, O.** (2026). Effectiveness of speed cameras in reducing speed: a systematic review. *Accident Analysis & Prevention*, 231, 108488. ISSN 0001-4575. <https://doi.org/10.1016/j.aap.2026.108488>
- Aristote.** *Éthique à Nicomaque*, livre V. (L'équité, *epieikeia*, comme correctif de la généralité de la loi.)
- Beccaria, C.** (1764). *Des délits et des peines*. (Fondements de la théorie de la dissuasion : certitude, sévérité, célérité de la sanction.)
- Bedsworth, F., Weber, B. & Willardsen, K.** (2025). Evaluating the effectiveness of freeway speed cameras: Evidence from a natural experiment in Dayton, Ohio. *Transportation Research Part A*, 200, 104649. doi:10.1016/j.tra.2025.104649.
- Bentham, J.** (1791). *Panopticon; or, The Inspection-House*. (Le principe panoptique : obtenir la conformité par la conscience d'une surveillance possible.)
- Blais, É. & Carnis, L.** (2015). Improving the safety effect of speed camera programs through innovations: Evidence from the French experience. *Journal of Safety Research*, 55, 135–145. doi:10.1016/j.jsr.2015.08.007.
- Callaway, B. & Sant'Anna, P. H. C.** (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- Carnis, L. & Blais, É.** (2013). An assessment of the safety effects of the French speed camera program. *Accident Analysis & Prevention*, 51, 301–309.
- Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P. C. & Sculpher, M.** (2015). Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technology Assessment*, 19(14). doi:10.3310/hta19140. Cité au §4.2.
- Cohn, E. G., Kakar, S., Perkins, C., Steinbach, R. & Edwards, P.** (2020). Red light camera interventions for reducing traffic violations and traffic crashes: A systematic review. *Campbell Systematic Reviews*, 16(2), e1091. doi:10.1002/cl2.1091.
- de Chaisemartin, C. & D'Haultfœuille, X.** (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.
- Duval, S. & Tweedie, R.** (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463.
- Dušek, L. & Traxler, C.** (2026). Swiftness and delay of punishment. *The Economic Journal* (article en ligne), doi:10.1093/ej/ueag035.
- Elvik, R.** (2005). Speed and road safety: Synthesis of evidence from evaluation studies. *Transportation Research Record*, 1908, 59–69.
- Elvik, R., Vadeby, A., Hels, T. & van Schagen, I.** (2019). Updated estimates of the relationship between speed and road safety at the aggregate and individual levels. *Accident Analysis & Prevention*, 123, 114–122.

- Foucault, M.** (1975). *Surveiller et punir. Naissance de la prison*. Paris : Gallimard. (Le panoptique : obtenir la conformité par la conscience d'une observation possible.)
- Gains, A., Nordstrom, M., Heydecker, B. & Shrewsbury, J.** (2005). *The national safety camera programme. Four-year evaluation report*. Londres : PA Consulting Group et University College London, décembre 2005, 160 p. Rapport commandité par le Department for Transport ; annexe H (« Estimates of regression-to-mean effects at safety cameras ») rédigée par L. Mountain (université de Liverpool) et M. Maher (Napier University). Périmètre : 38 partenariats, plus de 4 100 sites, avril 2000 à mars 2004. Sans correction de la RTM : -22 % de collisions corporelles (IC 95 % : -20 % à -24 %) et -42 % de tués ou blessés graves (IC 95 % : -40 % à -45 %). Après correction de tendance et correction *Empirical Bayes* sur un sous-échantillon de 216 sites urbains (chapitre 4 et annexe H), l'effet revient à -19 % pour les collisions corporelles comme pour les collisions graves ou mortelles, la RTM expliquant un quart de la baisse observée des collisions corporelles et trois cinquièmes de celle des collisions graves. Rapport bénéfiques sur coûts d'environ 2,7 pour 1, établi sur les seules collisions corporelles (résumé p. 2 ; sections 6.4 et 6.5 du rapport britannique, p. 82-83 ; note 4, p. 7). Le chapitre 5 mesure l'opinion des conducteurs (cf. §4.1). Rapports antérieurs de la même série : *A cost recovery system for speed and red-light cameras, two year pilot evaluation*, DfT, 11 février 2003 ; *The national safety camera programme, three-year evaluation report*, DfT, juin 2004.
- Gallagher, J. & Fisher, P. J.** (2020). Criminal deterrence when there are offsetting risks: Traffic cameras, vehicular accidents, and public safety. *American Economic Journal: Economic Policy*, 12(3), 202-237. doi:10.1257/pol.20170674. Cité au §11.2.
- Gallagher, J. & Fisher, P. J.** (2024). *Traffic Safety Program Evaluation: The Empirical Bayes Model and Mean Reversion Bias*. Transportation Research Board, 103^e Annual Meeting, communication TRBAM-24-02723. **Document de travail, non publié et non revu par les pairs.**
- Gao, J., Yang, D., Xu, C., Ozbay, K. & Sharma, S.** (2025). Assessing the impact of fixed speed cameras on speeding behavior and crashes: A longitudinal study in New York City. *Transportation Research Interdisciplinary Perspectives*, 30, 101373. doi:10.1016/j.trip.2025.101373. Plus de 1 800 radars fixes de zones scolaires ; baisse d'environ 14 % des accidents et de 75 % des infractions de vitesse à la fin de 2021 ; effet acquis en six mois. Cité au §3.5.
- Goodman-Bacon, A.** (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254-277.
- Gorell, R. & Sexton, B.** (2004). *Performance of safety cameras in London: Final Report*. Published Project Report PPR027, TRL Limited, décembre 2004 (copyright novembre 2004), ISBN 1-84608-028-2, rapport préparé pour Transport for London et le London Safety Camera Partnership. Étude avant/après de 77 sites de radars (au moins quatre collisions mortelles ou graves par kilomètre en trois ans) comparés à des sites témoins appariés, avec analyse complémentaire par Empirical Bayes ; réductions estimées de 12,4 % des collisions corporelles et de 20,6 % des collisions mortelles ou graves. <https://www.trl.co.uk/publications/ppr027>
- Graham, D. J., Naik, C., McCoy, E. J. & Li, H.** (2019). Do speed cameras reduce road traffic collisions? *PLOS ONE*, 14(9), e0221267. (Préprint : *Quantifying the causal effect of speed cameras on road traffic accidents via an approximate Bayesian doubly robust estimator*, arXiv:1703.05926.)
- Hauer, E.** (1997). *Observational Before-After Studies in Road Safety: Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*. Oxford : Pergamon / Elsevier Science.

- Hauer, E., Harwood, D. W., Council, F. M. & Griffith, M. S.** (2002). Estimating Safety by the Empirical Bayes Method: A Tutorial. *Transportation Research Record*, 1784, 126–131. doi:10.3141/1784-16.
- Høye, A.** (2014). Speed cameras, section control, and kangaroo jumps – a meta-analysis. *Accident Analysis & Prevention*, 73, 200–208. (63 estimations issues de 15 études sur radars fixes ; $\approx -20\%$ d'accidents, effet décroissant avec la distance ; -51% sur les accidents mortels, seul résultat dont l'auteure signale qu'il peut être affecté par le retour à la moyenne ; tests explicites de RTM, de biais de publication et de valeurs aberrantes.)
- Høye, A.** (2015). Safety effects of fixed speed cameras – An empirical Bayes evaluation. *Accident Analysis & Prevention*, 82, 263–269.
- ITF / OCDE** (2018). *Speed and Crash Risk*. International Transport Forum Research Report. Paris : OCDE.
- Job, R. F. S.** (2022). Evaluations of Speed Camera Interventions Can Deliver a Wide Range of Outcomes: Causes and Policy Implications. *Sustainability*, 14(3), 1765. doi:10.3390/su14031765.
- Juvénal.** *Satires*, VI, v. 347-348. (« Quis custodiet ipsos custodes ? », qui gardera les gardiens eux-mêmes ?)
- Nilsson, G.** (2004). *Traffic Safety Dimensions and the Power Model to Describe the Effect of Speed on Safety*. Bulletin 221, Lund Institute of Technology.
- Park, P. Y. & Lord, D.** (2010). Investigating Regression to the Mean in Before-and-After Speed Data Analysis. *Transportation Research Record*, 2165, 52–58. doi:10.3141/2165-06. (Distinction biais de sélection / RTM.)
- Persaud, B. & Lyon, C.** (2007). Empirical Bayes before-after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis & Prevention*, 39(3), 546–555.
- Pilkington, P. & Kinra, S.** (2005). Effectiveness of speed cameras in preventing road traffic collisions and related casualties: Systematic review. *BMJ*, 330(7487), 331–334.
- Roux, S. & Zamora, P.** (2013). L'impact local des radars fixes sur les accidents de la route : un effet important après l'installation mais plus réduit à long terme. *Économie et Statistique*, 460-461, novembre 2013, p. 37–68 (INSEE ; disponible sur Persée).
- Spiegelhalter, D.** (2013). Speed cameras, regression-to-the-mean, and the Daily Mail (again). Blog *Understanding Uncertainty*, 7 juin 2013. <https://understandinguncertainty.org/speed-cameras-regression-mean-and-daily-mail-again>
- Stagoff-Belfort, A., Ben-Menachem, J. & Beck, B.** (2025). Can speed cameras make streets safer? Quasi-experimental evidence from New York City. *PNAS*, 122(50), e2520328122. doi:10.1073/pnas.2520328122.
- Tyler, T. R.** (1990). *Why People Obey the Law*. New Haven : Yale University Press. (Théorie de la justice procédurale : légitimité perçue, « voix », obéissance volontaire.)
- Wilson, C., Willis, C., Hendrikz, J. K., Le Brocq, R. & Bellamy, N.** (2010). Speed cameras for the prevention of road traffic injuries and deaths. *Cochrane Database of Systematic Reviews*, 11, CD004607.

12.3 Sources institutionnelles et données

CEREMA (2014). *Performance des radars fixes dans les Yvelines*, V01, décembre 2014, affaire n° C14PM0072, 23 p. Cerema, direction territoriale Île-de-France, département Mobilité, unité Sécurité des déplacements et partage de la voirie ; rédacteurs U. Frontin et J.-F. Durand ; maître d'ouvrage DRIEA Île-de-France (SST/DSECR). Objet : adapter le parc des 35 radars fixes du département. Seul un extrait (p. 1 à 4 sur 23) est public : https://www.cerema.fr/fr/system/files?file=documents/2017/06/Etude2014_radars_fixes_78%20Extrait.pdf Trois précautions d'usage. La phase publiée classe les radars sur une période d'analyse de deux ans avant et après la mise en service (section 1.1.1 du rapport CEREMA) ; le protocole à cinq ans et deux zones figure séparément en annexe 5.3, intitulée « Méthode d'évaluation DSCR sur 5 ans » (sous-sections 5.3.1 à 5.3.4, zones de 1 000 m et de 4 000 m). La couverture porte « Décembre 2014 » et le pied de page « Décembre 2015 ».

Ifop pour Dimanche Ouest France (2011). *L'adhésion aux nouvelles mesures de sécurité routière. Résultats détaillés*, étude n° 110010, décembre 2011 (échantillon de 1 002 personnes de 18 ans et plus, méthode des quotas, terrain du 6 au 8 décembre 2011), avec rappel des enquêtes Ifop d'octobre 2003 (86 % d'approbation des premiers radars) et d'octobre 2004 (72 %). Citée au §2.7 et en note¹⁰. <https://www.ifop.com/publication/ladhesion-aux-nouvelles-mesures-de-securite-routiere/>

IGA, CGEDD, IGPN, IGGN (2014). *Évaluation de la politique de sécurité routière. Rapport de diagnostic*, tome 1, juillet 2014, 49 p. (IGA n° 14-074/14-016/01, CGEDD n° 009460-01, IGPN n° 13-1360-I, IGGN n° 1951 IGGN/CAB) ; tome 2, annexes 1 à 17, 320 p. Responsable opérationnelle : M. Bondaz, inspectrice générale de l'administration ; lettre de mission du ministre de l'Intérieur du 2 janvier 2014, faisant suite au CIMAP de décembre 2012. Formule citée au §3.6 : « Le contrôle sanction automatisé des vitesses a une efficacité avérée mais ne peut s'appliquer à tous les types d'infraction » (synthèse p. I ; section 3.2.1 du rapport IGA, p. 20). Le rapport n'avance aucune attribution chiffrée de la baisse de la mortalité au contrôle automatisé et relève que « la relation de cause à effet de chaque mesure n'est pas parfaitement établie » (p. 19-20). <https://www.onisr.securite-routiere.gouv.fr/sites/default/files/2025-05/14074-14016-01%20-%20MAP%20-%20Evaluation%20s%C3%A9curit%C3%A9%20routi%C3%A8re%20-%20rapport%20d%C3%A9finitif.pdf>

Cour des comptes (2021). *Évaluation de la politique publique de sécurité routière*, rapport public thématique, juin 2021 (mis en ligne le 1er juillet 2021), période évaluée 2008-2019, 17 recommandations ; référencé au catalogue sous l'intitulé « La politique publique de sécurité routière ». Le contrôle sanction automatisé y est crédité de « résultats incontestables, à un coût maîtrisé » (chapitre 1), assortis d'une demande de clarification de sa doctrine d'emploi et des objectifs d'implantation (recommandations 5 et 6, p. 24). Le rapport précise que « la Cour a fait le choix de ne pas évaluer le coût/avantage de la politique, en raison de l'insuffisance des données disponibles » (p. 14 ; p. 7 de la synthèse officielle). L'attribution de 75 % de la baisse de la mortalité au contrôle des vitesses ne figure nulle part dans le rapport complet (304 p.). Son tableau n° 1 (p. 38, d'après les rapports annuels de performance) donne l'évolution du parc par type de 2008 à 2019 : 1 473 radars fixes en 2008 sur 2 300 dispositifs, 4 097 dispositifs en 2013, 4 428 en 2018, 4 094 en 2019 (dont 375 tourelles). <https://www.ccomptes.fr/sites/default/files/2021-06/20210701-rapport-securite-routiere.pdf>

Cour des comptes (2023-2024). *Notes d'exécution budgétaire, Contrôle de la circulation et du stationnement routiers* (CAS CCSR) : lisibilité et affectation des recettes ; coûts de maintenance.

Direction du budget. *Utilisation par l'Agence de financement des infrastructures de transport de France (AFIT France) et par les collectivités territoriales des recettes issues du compte d'affectation spéciale « Contrôle de la circulation et du stationnement routiers », annexe générale au projet de loi de finances (document dit « jaune » budgétaire), éditions annuelles ; base légale : article 160 de la loi de finances pour 2017. Édition PLF 2025 : https://www2.assemblee-nationale.fr/static/17/Annexes-DL/PLF2025-Jaunes/17-Jaune_recettes-CAS-Radar.pdf Ce « jaune », le projet annuel de performances (« bleu », entrée suivante) et les notes d'exécution budgétaire de la Cour des comptes sont trois documents distincts, d'auteurs et de statuts différents.*

Délégation à la sécurité routière (2024). *Projet annuel de performances. Annexe au projet de loi de finances pour 2025. Compte d'affectation spéciale, mission interministérielle Contrôle de la circulation et du stationnement routiers, 69 p. (document dit « bleu » budgétaire), programme 751, présentation stratégique, p. 18 : « La vitesse excessive ou inadaptée constituant la première cause d'accident, la politique de déploiement des radars automatiques, décidée en 2002, a contribué à près des trois quarts de la baisse de la mortalité constatée entre novembre 2003 et décembre 2010. » Le document ne référence aucune étude à l'appui de cette phrase, reconduite d'une édition à l'autre. Il détaille par ailleurs le parc cible par type d'équipement (p. 28 à 31) : 4 160 équipements opérationnels et visibles, tous types confondus (300 radars fixes d'ancienne génération, 700 discriminants, 30 de vitesse moyenne, 300 feux rouges d'ancienne génération, 30 passages à niveau, 500 urbains, 1 500 tourelles, 550 autonomes de chantier, 150 voitures radar, 100 mobiles embarqués), hors environ 880 radars pédagogiques ; c'est la cible reprise au §2.2. Il donne aussi des enveloppes agrégées de maintien en condition opérationnelle, de déploiement et de pilotage, sans aucun coût unitaire d'installation. https://www2.assemblee-nationale.fr/static/17/Annexes-DL/PLF-2025/Controle_circulation_routiers.pdf*

Mission interministérielle « Contrôle et sanction automatisés des infractions au code de la route » (2009). *Rapports annuels de performances, annexe au projet de loi de règlement des comptes et rapport de gestion pour 2008, programme 751 « Radars », p. 35 (parc fin 2008 : 1 473 dispositifs fixes, 827 mobiles, 2 300 au total ; source du tableau n° 1 de la Cour des comptes 2021 et de la note³) et programme 752, bilan stratégique, p. 42 (« dans les zones accidentogènes couvertes par les radars, les accidents ont quasiment disparu », cité au §3.6). https://www.budget.gouv.fr/sites/performance_publique/files/farandle/ressources/2008/rap/pdf/DRGNORMALMSNYE.pdf*

Arrêté du 4 mars 2013 relatif à la signalisation d'une zone où la vitesse est contrôlée par un ou plusieurs dispositifs automatiques, JORF du 13 mars 2013 (modifie l'arrêté du 24 novembre 1967 relatif à la signalisation des routes et des autoroutes ; instruction interministérielle sur la signalisation routière du 7 juin 1977) : la signalisation « peut être effectuée » au moyen des panneaux SR3a ou SR3b (§7.1). <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000027167215>

Sénat (2017). V. Delahaye, *Sur la politique d'implantation des radars*, rapport d'information n° 644 (2016-2017), fait au nom de la commission des finances, déposé le 18 juillet 2017 (traçabilité et affectation des recettes ; qualifications de « tirelires » / « pompes à fric »). Le rapport cite la circulaire du 3 février 2004, qui prescrit de renseigner l'accidentologie de chaque site « sur section localisée de 850 mètres et sur une section plus longue de 4 000 mètres », et publie des coûts unitaires d'entretien allant d'environ 6 600 € par appareil pour un radar fixe double sens à environ 36 000 € pour un radar autonome de chantier, pour une durée de vie généralement retenue de dix ans. Il indique n'avoir obtenu ni coûts unitaires d'installation, ni série annuelle du parc. Il décrit en outre la méthodologie d'évaluation du CEREMA telle que la DSCR la lui a présentée (source indiquée : « réponse de la délégation à la sécurité et à la circulation routières au questionnaire du rapporteur spécial ») : deux zones, « l'une de 1 000 mètres

traduisant les comportements à proximité immédiate du radar, l'autre de 4 000 mètres englobant plus largement sa zone d'influence présumée », comparaison des données « sur des périodes de 5 ans avant et après l'installation en neutralisant l'année de pose », et un effet présenté comme ayant « permis, en termes de densité de tués, de ramener la zone de proximité immédiate du radar (1 000 mètres) au niveau moyen d'une zone plus large (4 000 mètres) » (§3.6). https://www.senat.fr/rap/r16-644/r16-644_mono.html

Assemblée nationale, question écrite n° 3437 de M. Olivier Falorni, « Impact des radars sur le taux de mortalité routière », publiée au JO AN du 22 novembre 2022, p. 5509 ; réponse du ministère de l'Intérieur et des Outre-mer publiée au JO AN du 21 mars 2023, p. 2703 : « une évaluation réalisée en 2013 a conclu que 75 % de la baisse de la mortalité entre 2003 et 2010 serait due au contrôle des vitesses par radar ». Même formulation dans la réponse à la question écrite n° 2852 (16^e législature). Ces réponses ne nomment pas l'évaluation invoquée : ni auteur, ni titre, ni support ; nous n'avons pas pu l'identifier. <https://questions.assemblee-nationale.fr/q16/16-3437QE.htm>

« **40 millions d'automobilistes** » : prises de position associatives de contestation (citées à titre de contexte du débat).

ANTAI (Agence nationale de traitement automatisé des infractions) ; système **LAPI** (lecture automatisée des plaques) ; **CNIL / RGPD** : cadre de traitement et de conservation des données du contrôle automatisé (cf. §4.3).

Sondages d'opinion sur la perception « recettes contre sécurité » du contrôle automatisé, cf. §2.7 et §4.1. France : sondage BVA réalisé par téléphone du 12 au 23 mai 2014 auprès d'un échantillon de conducteurs, commandé dans le cadre de l'évaluation interministérielle et annexé au rapport IGA/CGEDD/IGPN/IGGN de juillet 2014 (tome 1, p. 16 ; annexe 15 du tome 2) : « 79 % des conducteurs considèrent que les actions menées aujourd'hui en France par les pouvoirs publics en matière de sécurité routière ont pour objectif de remplir les caisses de l'État (dont 43 % “tout à fait”) » et près de 8 conducteurs sur 10 estiment que les radars sont placés dans des endroits qui rapportent de l'argent plus que dans des endroits dangereux. Le même sondage donne, en sens inverse, 84 % de conducteurs jugeant encore possible de diminuer largement le nombre de tués, 72 % estimant que les actions vont dans le bon sens, 71 % qu'elles réduisent réellement le nombre de victimes et 62 % qu'elles reposent sur des statistiques fiables (p. 17) ; le rapport qualifie la première série de perceptions de « parfaitement faux » et oppose environ 0,8 Md€ de recettes de radars à environ 4 Md€ de dépenses de l'État pour la sécurité routière. Royaume-Uni : Gains et al. (2005), chapitre 5 (§12.2). Aucune enquête états-unienne mesurant cette perception n'a pu être vérifiée sur son document d'origine ; les chiffres correspondants ont donc été écartés.

ONISR / BAAC : fichiers d'accidents corporels, data.gouv.fr, millésimes 2017 à 2024 pour l'ensemble de l'étude et millésimes 2005 à 2016 pour la seule mesure de couverture des coordonnées citée au §5. Les estimations n'utilisent que caractéristiques (date, coordonnées) et usagers (gravité). Le champ agg de caractéristiques (milieu) et le champ catr des fichiers lieux (catégorie de route) servent à la comparaison descriptive des deux groupes publiée à l'annexe 13.6, section D, et à la mesure de complétude du géocodage citée au §11 ; ils n'entrent dans aucune estimation ; les fichiers vehicules n'entrent dans aucun calcul publié.

Fichier des radars fixes : jeu de données des radars automatiques publié sur data.gouv.fr (numéro, type, date de mise en service, coordonnées, vitesse maximale autorisée), sous Licence Ouverte Etalab 2.0. Ce jeu est mis à jour en continu : l'extraction utilisée par l'étude compte 3 268 enregistrements, dont 3 167

en France métropolitaine et 701 sans numéro de radar ; elle est conservée avec le code, comme l'indiquent les prérequis de `pipeline_reproduction.md`. Une extraction antérieure du même jeu (état interne du 5 septembre 2018 : 3 350 enregistrements, 3 287 en métropole, avec l'équipementier, la direction contrôlée et la date d'installation), téléchargée en août 2026 depuis dev.data.gouv.fr/datasets/radars-automatiques, sert à qualifier la nature du traitement (§5) ; elle est conservée avec le code au même titre.

SDES, *Bilan de la circulation* (Service des données et études statistiques, ministère chargé des transports) : tableau « G1.c Circulation par réseau », en milliards de véhicules-kilomètres. Ligne « Ensemble des réseaux » pour la variante de tendance trafic (annexe 13.4) ; évolutions par réseau pour les valeurs citées au §11. L'ensemble des séries constitutives du bilan a été rebasé en 2023, d'où une rupture visible entre 2022 et 2023.

ONISR, *Description des bases de données annuelles*, octobre 2025 : dictionnaire des variables des fichiers BAAC, dont la variable de gravité et la définition du tué à trente jours ; source de la vérification d'absence de rupture de série sur la mortalité 2017-2024 (§5).

Didou, O. & Limousin, F. (2025). *Immatriculations des voitures particulières en 2024 : baisse dans le neuf et léger rebond pour l'occasion*, SDES, 24 février 2025, données RSVERO (1,76 million d'immatriculations de voitures neuves en 2024) ; **Didou, O.** (2024). *39,3 millions de voitures en circulation en France au 1er janvier 2024*, SDES, 23 septembre 2024, données RSVERO provisoires (âge moyen 11,2 ans ; valeurs révisées à 39,4 millions et 11,1 ans dans la publication du 19 septembre 2025, d'où l'arrondi retenu au §11), <https://www.statistiques.developpement-durable.gouv.fr/393-millions-de-voitures-en-circulation-en-france-au-1er-janvier-2024> Ces valeurs servent uniquement à calculer l'ordre de grandeur du rythme de renouvellement du parc (environ 4,5 % par an) invoqué au §11 ; elles n'entrent dans aucune régression et ne sont reproduites par aucune étape de nos calculs.

Règlement (UE) 2019/2144 du Parlement européen et du Conseil du 27 novembre 2019 relatif aux prescriptions applicables à la réception par type des véhicules à moteur et de leurs remorques, ainsi que des systèmes, composants et entités techniques distinctes destinés à ces véhicules, en ce qui concerne leur sécurité générale et la protection des occupants des véhicules et des usagers vulnérables de la route (dit « GSR2 »). Cité au §11 pour ses dates d'application : nouveaux types de véhicules à compter du 6 juillet 2022, tous véhicules neufs à compter du 7 juillet 2024, d'où sa portée négligeable sur la fenêtre d'observation 2017-2024. <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32019R2144>

Sources et scripts de l'étude : realitesroutieres.fr/etude-radars-efficacite-2026.

13. Annexes

13.1 Sorties de régression : fréquence des accidents

Régression de Poisson (erreurs-types regroupées par site, *cluster*), structure « tendance + cassure ». La ligne post porte l'effet radar (rupture à l'installation).

```

=====
RAYON 500 m - NB_ACC
=====
Poisson Regression Results
=====
Dep. Variable:          y      No. Observations:          17206
Model:                Poisson  Df Residuals:          17195
Method:                MLE     Df Model:              10
                             Pseudo R-squ.:          0.06120
                             Log-Likelihood:         -72053.
converged:              True   LL-Null:             -76750.
Covariance Type:        cluster LLR p-value:           0.000
=====

```

	coef	std err	z	P> z	[0.025	0.975]
Intercept	1.2745	0.077	16.597	0.000	1.124	1.425
C(an) [T.2018]	-0.0053	0.025	-0.214	0.831	-0.054	0.043
C(an) [T.2019]	0.0745	0.031	2.417	0.016	0.014	0.135
C(an) [T.2020]	-0.1126	0.033	-3.435	0.001	-0.177	-0.048
C(an) [T.2021]	0.0368	0.037	0.998	0.318	-0.035	0.109
C(an) [T.2022]	-0.0407	0.037	-1.110	0.267	-0.112	0.031
C(an) [T.2023]	-0.0486	0.041	-1.181	0.237	-0.129	0.032
C(an) [T.2024]	-0.0653	0.045	-1.462	0.144	-0.153	0.022
traite	-1.1317	0.108	-10.512	0.000	-1.343	-0.921
evt_lin	-0.0737	0.018	-3.984	0.000	-0.110	-0.037
post	0.0338	0.053	0.642	0.521	-0.069	0.137

```

=====

```

Lecture. Effet radar, ligne post : +3,4 %, p = 0,521, IC 95 % [-6,7 % ; +14,7 %] (conversion du coefficient post = 0,0338 et de son intervalle ; valeurs produites telles quelles par les étapes 5 et 11 du programme de reproduction).

```

=====
RAYON 1000 m - NB_ACC
=====
Poisson Regression Results
=====
Dep. Variable:          y      No. Observations:          20790
Model:                  Poisson  Df Residuals:              20779
Method:                  MLE     Df Model:                  10
                               Pseudo R-squ.:              0.07253
                               Log-Likelihood:             -1.7539e+05
converged:              True    LL-Null:                -1.8910e+05
Covariance Type:        cluster LLR p-value:                0.000
=====
               coef      std err          z      P>|z|      [0.025      0.975]
-----
Intercept          2.0361      0.078      26.097      0.000       1.883       2.189
C(an) [T.2018]     -0.0014      0.021      -0.069      0.945      -0.042       0.039
C(an) [T.2019]       0.0647      0.029       2.254      0.024       0.008       0.121
C(an) [T.2020]     -0.1110      0.033      -3.409      0.001      -0.175      -0.047
C(an) [T.2021]       0.0576      0.037       1.548      0.122      -0.015       0.131
C(an) [T.2022]     -0.0365      0.036      -1.009      0.313      -0.107       0.034
C(an) [T.2023]     -0.0213      0.040      -0.531      0.595      -0.100       0.057
C(an) [T.2024]     -0.0372      0.044      -0.852      0.394      -0.123       0.048
traite             -1.2108      0.110     -10.981      0.000      -1.427      -0.995
evt_lin            -0.0894      0.018      -4.987      0.000      -0.125      -0.054
post                0.0475      0.043       1.105      0.269      -0.037       0.132
=====

```

Lecture. Effet radar, ligne post : +4,9 %, $p = 0,269$, IC 95 % [-3,6 % ; +14,1 %] (conversion du coefficient post = 0,0475 et de son intervalle ; étapes 5 et 11).

Repères complémentaires (rayon 500 m). Le DiD naïf, sans le terme de pente `evt_lin`, donne **-28,9 %** ($p < 0,001$, IC [-40 % ; -16 %]) : c'est l'ampleur du biais de retour à la moyenne qu'absorbe la pente. L'étude d'événement (référence = année -1, année d'installation exclue) situe les périodes anciennes au-dessus de la référence (trois ans et plus avant la pose : +0,18, $p = 0,004$; deux ans avant : +0,02, $p = 0,72$) et les périodes postérieures en dessous (+1 an : -0,12, $p = 0,022$; deux ans et plus après : -0,32, $p < 0,001$).

Cette étude d'événement est estimée en coupe **poolée** : les temps relatifs éloignés ne sont renseignés que par les cohortes d'installation tardives (voir les effectifs par temps relatif en annexe 13.6, section B). Les coefficients mélangent donc un effet de temps relatif et un effet de composition en cohortes, et ne s'interprètent pas directement comme une trajectoire. L'écart est important : le coefficient à -5 ans vaut +0,39 dans l'étude d'événement poolée restreinte aux années antérieures à la pose, et -0,14 à l'intérieur des cohortes, sur exactement les mêmes 414 zones. Estimé à composition constante sur les cinq années qui

précèdent la pose, le profil antérieur à la pose est **croissant** et culmine l'année précédant l'installation (annexe 13.7) : c'est la signature d'une sélection réactive sur un pic, mais non d'un reflux déjà engagé plusieurs années avant la pose, comme la lecture poolée le suggérerait.

13.2 Résultats de gravité (fenêtre homogène 2018-2024)

Format de lecture : indicateur | rayon | événements avant→après | cassure (effet) p IC | seuil détectable.

=== GRAVITÉ, fenêtre homogène 2018-2024 ===									
nb_tues	500m	événements	230→ 134	cassure	-11%	p=0.619	IC[-44;+41]		seuil détectable 48%
nb_tues	1000m	événements	497→ 344	cassure	-6%	p=0.712	IC[-31;+29]		seuil détectable 36%
nb_hospit	500m	événements	1385→ 771	cassure	-2%	p=0.886	IC[-20;+21]		seuil détectable 25%
nb_hospit	1000m	événements	3334→1937	cassure	+9%	p=0.257	IC[-6;+25]		seuil détectable 18%
nb_ksi	500m	événements	1615→ 905	cassure	-3%	p=0.766	IC[-20;+18]		seuil détectable 24%
nb_ksi	1000m	événements	3831→2281	cassure	+7%	p=0.356	IC[-7;+22]		seuil détectable 18%

Interprétation (exemple : nb_tues | 500m). nb_tues désigne les tués (nb_hospit = blessés graves ; nb_ksi = les deux réunis). 500m : on compte ce qui se passe dans 500 m autour du radar. 230→134 : 230 tués avant l'installation, 134 après, dans ces zones. cassure -11 % : l'effet attribué au radar est une baisse de 11 % ; **ce n'est pas** le simple rapport 230/134, mais la baisse qui subsiste une fois retirée la tendance naturelle de baisse de la mortalité. p = 0,619 et IC[-44 ; +41] : valeur de p et intervalle de confiance à 95 %, dont la définition est donnée au §8. L'intervalle contient ici une baisse *et* une hausse, donc on ne peut pas conclure. seuil détectable 48 % : la plus petite baisse repérable de façon fiable est de 48 % ; en deçà, l'effet reste invisible faute de données suffisantes sur les décès.

Ce tableau confirme, sur la fenêtre homogène 2018-2024, l'absence de signal sur les tués à 500 m (-11 %, p = 0,62), cohérente avec la fenêtre complète 2017-2024 (-17 %, p = 0,38, §8).

13.3 Méthodologie statistique

Données. Traités = zones dans un rayon R (500 m / 1 km) des radars installés 2019-2024 ; témoins = zones dans R des radars installés avant 2019 (pas de changement sur la période). Chaque accident n'est rattaché qu'au radar le plus proche : la zone n'est donc pas le disque de rayon R , mais ce disque tronqué à mi-distance des radars voisins (§11.3). Une ligne = zone $\times$ année ; réponse = compte d'issues (nb_tues, nb_hospit, ou leur somme nb_ksi) ; offset = couverture GPS. Fenêtres : 2017-2024 pour la fréquence et pour les tués ; 2018-2024 pour les analyses de gravité (blessés graves, KSI, et sensibilité sur les tués), 2017 étant exclu en raison de la rupture de saisie de grav=3. L'année d'installation de chaque zone traitée est exclue car transitoire.

Spécification. Régression de Poisson, structure « tendance + cassure » :

$$\log E[Y_{zt}] = \log(c_t) + \beta_0 + \beta_1 \cdot T_{zt} + \sum_y \gamma_y \cdot 1\{an=y\} + \delta \cdot k_{zt} + \theta \cdot post_{zt}$$

- T_{zt} : indicatrice zone traitée (décalage de niveau entre groupes).
- γ_y : effets fixes d'année (tendance commune, identifiée principalement par les témoins).
- k_{zt} (ligne evt_lin des sorties de régression de l'annexe 13.1) : temps relatif à l'installation ($an - \text{année d'install}$) pour les traités, 0 pour les témoins $\rightarrow$ pente continue, destinée à absorber la tendance propre aux traités, retour à la moyenne compris ; sur ce que ce terme capte effectivement, voir plus bas et l'annexe 13.6, section E.
- $post_{zt}$: 1 si traité et $an > \text{année d'install}$ $\rightarrow$ indicatrice de traitement.
- c_t : couverture GPS (offset / exposition ; part nationale des accidents à coordonnées renseignées, 0,87 en 2017, 0,97 en 2018, 1 ensuite, voir §7.3).

θ est le paramètre d'intérêt : la rupture discrète à l'installation, nette de la pente continue et de la tendance calendaire. Effet = $\exp(\theta) - 1$. Le cœur de l'identification est la décomposition k_{zt} (pente) / $post_{zt}$ (saut) : θ mesure l'écart au prolongement de la pente, et non l'écart au niveau du pic. C'est ce qui distingue cette spécification d'un DiD naïf que le retour à la moyenne biaisait.

Portée exacte de cette identification. La pente k_{zt} est estimée **conjointement sur les observations antérieures et postérieures** à l'installation ; elle n'est pas estimée sur la seule période « avant » puis extrapolée. θ mesure donc l'écart au prolongement d'une tendance ajustée sur tout le support, sous l'hypothèse que le reflux se poursuit après la pose au même

rythme qu'avant. Cette hypothèse n'est pas testable et elle est déterminante : une variante estimant la pente sur les seules années antérieures donne, sur la fréquence à 500 m, $-16,1\%$ $[-25,0 ; -6,2]$ ($p = 0,002$) au lieu de $+3,4\%$. Cf. l'annexe 13.6 pour plus d'informations. Cette identification a une conséquence symétrique du côté de l'effet. θ est une discontinuité datée à la pose, non un effet moyen sur la période postérieure : la pente k_{zt} étant estimée conjointement sur l'avant et sur l'après, un effet qui s'installe progressivement (activation différée, apprentissage des conducteurs, signalisation posée après la cabine) ou qui s'estompe se reporte en partie sur cette pente et n'est pas restitué par θ . Deux conséquences en découlent. La comparaison avec les évaluations dont le résultat central est un effet variable dans le temps (§3.6) n'est pas une comparaison de même nature. Et un θ unique est imposé à des cohortes dont la durée d'observation postérieure va de zéro à cinq ans (les 111 zones de la cohorte 2024, sans année postérieure, restent dans l'estimation et contribuent au niveau et à la pente, non à θ ; annexe 13.4) : les éléments dynamiques disponibles sont les fenêtres événementielles $[-2 ; +2]$ et $[-3 ; +3]$ de l'annexe 13.6, section A, et les ATT(g, t) de l'annexe 13.5, dont la dispersion entre cohortes et années va de $-30,7\%$ à $+9,5\%$.

Estimation et inférence. Poisson par maximum de vraisemblance (statsmodels, method="bfgs", maxiter=5000). Erreurs-types robustes regroupées par zone (cluster sur le radar) : elles corrigent la corrélation sérielle intra-zone et la surdispersion ; sans regroupement, les erreurs-types de Poisson seraient sous-estimées et la significativité gonflée. Les « événements 230→134 » sont les comptes bruts avant/après (élément de lecture, pas l'estimand) ; l'effet vient de θ et de son intervalle de Wald.

Traitement du facteur confondant principal. La sélection réactive (radars posés après un pic) est traitée par le seul terme k_{zt} , pente propre aux traités ; ce terme n'est cependant pas identifié par la trajectoire de chaque site : en l'absence d'effet fixe de zone, et le temps relatif valant l'année moins le millésime de pose avec des indicatrices d'année saturées, il capte pour partie le gradient de niveau entre cohortes de pose (annexe 13.6, section E) ; les effets fixes d'année γ_y , identifiés principalement par les zones témoins, qui n'ont ni pente propre ni indicatrice de traitement, absorbent la tendance calendaire commune aux deux groupes et non la sélection (les zones traitées y contribuent également, par la part de leur variation calendaire que la pente en temps relatif ne capte pas). Ce traitement reste imparfait : le placebo de l'annexe 13.7 montre que la spécification ne restitue pas zéro là où l'effet vaut zéro, une pente monotone ne pouvant pas épouser une trajectoire qui monte puis redescend. L'hypothèse de tendances parallèles a été examinée par une étude d'événement préalable, qui révélait justement la violation en spécification naïve, d'où l'ajout de la pente.

Puissance. Effet minimal détectable à 80 % de puissance, test bilatéral, $\alpha = 0,05$: $|\theta_{\min}| = (z_{0,975} + z_{0,80}) \cdot SE(\theta) = 2,80 \cdot SE(\theta)$. L'échelle du modèle étant logarithmique, ce seuil est asymétrique : il vaut $1 - \exp(-|\theta_{\min}|)$ en réduction et $\exp(|\theta_{\min}|) - 1$ en hausse. La colonne « seuil détectable » du §8 reporte la borne à la baisse ; les deux bornes sont données au §9 et sont calculées aux étapes 5 et 11 du programme de reproduction. Ce seuil borne ce que l'échantillon peut voir.

Empirical Bayes (Hauer, « EB ») : borne basse et diagnostic, non une seconde correction. Modèle de référence binomial négatif ajusté sur la période avant **des sites traités eux-mêmes** $\rightarrow \mu$ (attendu) et α (surdispersion, $\text{Var} = \mu + \alpha \cdot \mu^2$). Ce choix s'écarte de la SPF canonique, calibrée sur des sites de référence non traités (§3.3.3) : l'écart de niveau entre traités et témoins (facteur $\approx 0,3$, §11) rendrait ce référentiel externe inadapté ici. La contrepartie est que la part de retour à la moyenne commune à l'ensemble du groupe traité est intégrée à μ et disparaît de l'estimation (Gallagher & Fisher, 2024) : avec un poids médian $w = 0,049$ sur la fréquence (0,493 sur les tués), le rétrécissement bayésien ne déplace l'estimation que de 2,5 points, et dans le sens d'une baisse plus forte (annexe 13.4). L'Empirical Bayes est donc ici une borne basse et un diagnostic de cohérence, non une correction de la RTM ; sa comparaison avec le test de cassure, qui identifie l'effet par une autre voie, sert de diagnostic et non de double correction. Ce modèle de référence est sans covariable (constante et exposition de couverture GPS seules) : l'attendu annuel $\bar{\mu}$ est donc commun à tous les sites. Pour un site donné, l'attendu de la période antérieure vaut $M = \bar{\mu} \times \Sigma c_t$ sur les années « avant » du site. Poids $w = 1/(1 + \alpha \cdot M)$; attendu corrigé $\hat{m} = w \cdot M + (1-w) \cdot K_{\text{avant}}$; projection vers l'après par le ratio de tendance du groupe témoin $\rho : \pi = \hat{m} \cdot \rho$; indice d'efficacité $\theta_{\text{EB}} = (K_{\text{après}}/\pi)/(1 + \text{Var}(\pi)/\pi^2)$; effet = $1 - \theta_{\text{EB}}$. L'intervalle et le test de Wald qui l'accompagnent traitent le ratio de tendance ρ et les deux paramètres du modèle de référence comme connus sans erreur : ils sont conditionnels à ces grandeurs, et d'autant plus étroits que l'issue projetée est rare, la tendance du groupe témoin étant alors mesurée sur peu d'événements. La concordance ou la divergence entre EB et cassure sert d'indicateur de fiabilité (concordance sur la fréquence, divergence sur les tués aux deux rayons : EB -30,1 % contre cassure -17 % à 500 m, EB -24,3 % contre cassure -7,5 % à 1 km).

LIMITES STATISTIQUES

- *TWFE en adoption échelonnée.* L'estimateur à effets fixes bidirectionnels peut être biaisé sous hétérogénéité des effets de traitement, via les « comparaisons interdites » entre cohortes (Goodman-Bacon, 2021 ; de Chaisemartin & D'Haultfœuille, 2020 ; Callaway & Sant'Anna, 2021). L'approche EB par site avant/après atténue ce biais ; une extension rigoureuse emploierait un estimateur robuste à l'hétérogénéité (par ex. DiD de Callaway-

Sant'Anna avec témoins « not-yet-treated »). Ce contrôle a été mis en œuvre sur la fréquence des accidents, avec deux groupes de comparaison (zones pas encore traitées, radars anciens) : il ne fait apparaître aucun effet, pour un effet minimal détectable de 14 à 17 % à la baisse, calculé sur l'échelle logarithmique à partir des intervalles bootstrap (annexe 13.5). Il n'a pas été étendu à la mortalité, où la rareté des événements rend l'estimation cohorte par cohorte trop instable.

- *Surdispersion* : gérée par les SE cluster ; une binomiale négative serait l'alternative paramétrique (c'est d'ailleurs ce que fait l'EB).
- *Événements rares* : pour les tués, faible « quantité », puissance intrinsèquement limitée (seuils détectables 33-48 % selon la fenêtre et le rayon), non corrigeable par le modèle.
- *Tendances parallèles* : hypothèse non testable directement ; les tendances pré-installation sont examinées par l'étude d'événement sur la fréquence (annexe 13.1), la puissance manquant pour cet examen sur les issues rares.
- *Régression de noyau (kernel regression)* : ne semble pas appropriée d'après des tests sur des rayons plus larges qu'1 km.

13.4 Paramètres détaillés de l'Empirical Bayes et tailles d'échantillon

Cette annexe publie les paramètres intermédiaires de l'estimation Empirical Bayes et les tailles d'échantillon de chaque analyse, afin de permettre la vérification indépendante des calculs.

Tailles d'échantillon des régressions de cassure. Une « zone » désigne l'abord d'une cabine radar du fichier des radars, identifiée par sa ligne dans le fichier (ce qui inclut les 701 cabines sans numéro, §7.3), ayant enregistré au moins un accident géolocalisé sur la fenêtre d'étude : sur les 1 521 radars installés entre 2019 et 2024, 1 002 sont le radar le plus proche d'au moins un accident survenu à moins de 500 m (1 290 à moins de 1 km) ; les zones sans aucun accident sur la période sont exclues par construction. L'année d'installation de chaque radar traité est exclue (année transitoire).

RAYON	FENÊTRE	ZONES TRAITÉES	ZONES TÉMOINS	LIGNES ZONE × ANNÉE
500 m	2017-2024	1 002	1 274	17 206
1 km	2017-2024	1 290	1 470	20 790
500 m	2018-2024 (gravité)	1 002	1 274	14 930
1 km	2018-2024 (gravité)	1 290	1 470	18 030

Convention d'univers. Le constat descriptif du §6 retient les zones ayant enregistré au moins un accident dans la fenêtre considérée (948 zones pour 2018-2024), tandis que les régressions retiennent l'univers défini sur 2017-2024 (1 002 zones, tableau ci-dessus). Les comptages d'événements sont identiques dans les deux cas, les zones sans accident dans la fenêtre contribuant zéro.

Paramètres de l'Empirical Bayes (fréquence des accidents). Le modèle de référence est une binomiale négative calée sur la période « avant » des sites traités, avec la couverture GPS en exposition. Les sites exploitables sont ceux disposant d'une exposition non nulle avant *et* après l'installation ; les radars installés en 2024, sans année complète « après », sortent de l'analyse (1 002 zones : 891 sites à 500 m). Convention de lecture : effet négatif = baisse.

PARAMÈTRE	500 M	1 KM
Sites exploitables	891	1 149
μ (accidents attendus / site / an)	1,432	2,970
α (surdispersion ; $\text{Var} = \mu + \alpha \cdot \mu^2$)	3,54	4,01
Poids w (médiane)	0,049	0,021
ρ (médiane)	0,72	0,72
Accidents observés après (λ)	2 679	6 948
Attendus sans radar (Π ; $\text{Var}(\Pi)$)	2 772,0 (2 863,5)	7 089,2 (7 424,7)
Effet EB	-3,4 % [-8,6 ; +1,8]	-2,0 % [-5,3 ; +1,3]
p (Wald sur θ)	0,20	0,23

Le comptage « après » ($\lambda = 2\,679$ à 500 m) coïncide avec celui du constat descriptif du §6 : les 111 zones écartées de l'Empirical Bayes (installations 2024) n'ont aucune année postérieure. Le comptage « avant » diffère en revanche (4 721 sur les 891 sites de l'EB, contre 5 869 sur les 1 002 zones du §6). La convention d'exclusion de l'année d'installation est identique partout.

Lecture. Un poids w proche de 0 signifie que l'estimation de ce qui se serait passé sans radar repose surtout sur l'historique propre du site, plutôt que sur le modèle de référence ; un poids proche de 1 signifie l'inverse. $\bar{\mu}$ est un attendu annuel, exprimé par unité de couverture GPS ; le poids w se calcule sur l'attendu de la période antérieure propre à chaque site ($\bar{\mu}$ multiplié par la somme des couvertures des années « avant » du site), et non sur $\bar{\mu}$ lui-même : c'est pourquoi le w médian publié est très inférieur à celui qu'on obtiendrait en appliquant la formule directement au taux annuel.

Implications. Avec un poids médian $w = 0,049$ sur la fréquence, le contrefactuel repose à environ 95 % sur l'historique propre du site, c'est-à-dire sur la grandeur même que la sélection réactive a gonflée. Le modèle de référence étant de surcroît calibré sur la période « avant » **des sites traités eux-mêmes** (faute de référentiel externe comparable, l'écart de niveau avec les témoins étant d'un facteur 3), la part de retour à la moyenne commune à l'ensemble du groupe traité est intégrée au contrefactuel et disparaît de l'estimation. Vérifications : en forçant $w = 0$ (aucun rétrécissement bayésien), l'effet vaut $-0,9\%$ contre $-3,4\%$ avec ; le seul ingrédient anti-RTM de la chaîne déplace donc l'estimation de 2,5 points, et dans le sens d'une baisse plus forte. Sur les tués (w médian 0,493), l'Empirical Bayes donne

–30,1 % là où l'avant/après non corrigé du §6 donne –31,0 % : il reproduit l'estimation non corrigée. **L'Empirical Bayes est ici un diagnostic de cohérence et une borne basse, non une correction indépendante de la RTM (§3.7).**

Décomposition de l'écart entre le brut et l'effet corrigé, sur les 891 sites retenus par l'Empirical Bayes (accidents, 500 m). Sur l'univers du §6 (1 002 zones, 5 869 accidents avant), l'avant/après brut vaut –29,9 % ; sur les 891 sites de l'Empirical Bayes (4 721 accidents avant, les installations 2024 étant écartées faute d'année postérieure), il vaut –31,8 %. C'est de ce dernier point de départ que part la décomposition, en trois étapes dont la contribution respective est la suivante :

ÉTAPE	CONTREFACTUEL	EFFET	ÉCART CUMULÉ
Avant/après poolé à exposition égale	3 926	–31,8 %	n.a.
Passage à une pondération site à site (rapport d'exposition propre)	3 014	–11,1 %	20,7 pts
Ajout de la tendance du groupe témoin ρ (équivalent à $w = 0$)	2 703	–0,9 %	10,2 pts
Ajout du rétrécissement bayésien (w médian 0,049)	2 772	–3,4 %	–2,5 pts

Environ deux tiers de l'écart proviennent donc du changement de pondération (la comparaison poolée sur-pondère les cohortes tardives dans le « avant » et les cohortes précoces dans le « après » ; la corrélation entre comptage « avant » et ρ vaut –0,20), et un tiers de la tendance du groupe témoin, c'est-à-dire de la maturation. Le retour à la moyenne proprement dit n'apparaît pas dans cette chaîne : il est identifié par le seul terme de pente du test de cassure ($k_{zt} = -0,0737$ par an, $p < 0,001$ à 500 m). Ce terme est compatible avec toute tendance propre aux zones traitées (retour à la moyenne, mais aussi renouvellement de chaussée, évolution du trafic local ou effet d'anticipation), et rien dans les données ne permet de les départager. Ainsi : *l'essentiel de la baisse brute n'est pas attribuable au radar*. Le ratio ρ intègre à la fois la tendance du groupe témoin, la durée d'observation relative de chaque site et, dans la variante principale, la couverture GPS : la série témoin y est un comptage géolocalisé brut, affecté par la sous-couverture de 2017 et 2018 au même titre que le comptage « avant » du site, de sorte que les deux se compensent. Le ratio varie de part et d'autre de 1 selon l'année d'installation (médiane inférieure à 1 : la majorité des sites sont observés moins longtemps après qu'avant). Il en va autrement de la variante à tendance trafic ci-dessous : la série de véhicules-kilomètres du SDES ne comporte aucun facteur de couverture, et ρ n'y compense pas la sous-couverture des deux premières années. Les deux variantes ne sont donc pas construites sur la même échelle, et l'écart entre –3,4 % et –7,5 %

n'est pas entièrement imputable au changement de contrefactuel. Tant qu'un contrôle déflatant la série de trafic par la couverture GPS n'a pas été conduit, le -7,5 % s'entend sous cette réserve.

Robustesse : tendance mesurée sur le trafic national (§7.3). En remplaçant la tendance du groupe témoin par le trafic national réel (SDES, creux Covid 2020 inclus), l'effet EB sur la fréquence passe de -3,4 % à **-7,5 %** [-12,3 ; -2,6] ($p = 0,003$) à 500 m et de -2,0 % à **-5,5 %** [-8,6 ; -2,4] ($p < 0,001$) à 1 km. L'écart s'explique par le choix du contrefactuel, c'est-à-dire ce qui se serait passé sans radar. La variante trafic suppose que, sans radar, les accidents auraient suivi le trafic national : elle attribue donc au radar toute évolution des abords de radars qui s'écarte du trafic. Or les zones témoins (radars anciens, sans changement sur la période) connaissent la même baisse relative au trafic (ρ médian 0,72 contre 0,76) : cette baisse, commune à tous les abords de radars, ne peut pas être un effet des installations récentes, et c'est précisément ce que neutralise le contrefactuel principal, construit sur le groupe témoin. À forme de tendance fixée, les estimations sur la fréquence restent proches : de -7,5 % (EB avec tendance trafic, 500 m) à +4,9 % (test de cassure à 1 km, tableau du §8), en passant par l'EB à tendance témoin (-3,4 % / -2,0 %) et le contrôle Callaway-Sant'Anna (variantes A et B à 500 m : -3,2 % et -0,7 %, annexe 13.5). Cette proximité ne vaut que pour le choix du contrefactuel : l'annexe 13.6 montre que faire varier la forme de la tendance déplace l'estimation bien davantage, et c'est cette dépendance qui fonde la réserve du §8.

Paramètres de l'Empirical Bayes (mortalité). Même structure que pour la fréquence, avec une différence de construction : la tendance du groupe témoin ρ est mesurée sur les **tués** des zones témoins (et non sur leur nombre d'accidents), par cohérence avec l'indicateur projeté. Les paramètres $\bar{\mu}$ et α sont estimés sur les tués de la période « avant » des sites traités.

PARAMÈTRE	500 M	1 KM
Sites exploitables	891	1 149
μ (tués attendus / site / an)	0,0729	0,1253
α (surdispersion)	3,67	2,89
Poids w (médiane)	0,493	0,418
ρ (médiane)	0,78	0,84
Tués observés après (λ)	134	344
Attendus sans radar (Π ; Var(Π))	191,3 (109,5)	453,8 (321,8)
Effet EB	-30,1 % [-44,1 ; -16,2]	-24,3 % [-34,2 ; -14,4]
p (Wald sur θ)	< 0,001	< 0,001

Paramètres de l'Empirical Bayes (gravité). Même structure, sur la fenêtre homogène 2018-2024 de la gravité (§5), la tendance du groupe témoin étant mesurée sur le même indicateur que l'issue projetée ; ces estimations complètent le diagnostic, d'abord limité à la fréquence et à la mortalité.

PARAMÈTRE	BLESSÉS GRAVES, 500 M	BLESSÉS GRAVES, 1 KM	KSI, 500 M	KSI, 1 KM
Sites exploitables	891	1 149	891	1 149
μ (attendus / site / an)	0,431	0,820	0,502	0,942
α (surdispersion)	2,76	2,79	2,32	2,51
Poids w (médiane)	0,221	0,128	0,224	0,125
Effet EB	-6,3 % [-15,4 ; +2,8]	-0,8 % [-7,2 ; +5,6]	-8,8 % [-17,0 ; -0,6]	-2,5 % [-8,4 ; +3,4]
p (Wald sur θ)	0,18	0,80	0,036	0,40
Test de cassure (§8, même fenêtre)	-1,5 % (p = 0,89)	+8,6 % (p = 0,26)	-2,9 % (p = 0,77)	+6,7 % (p = 0,36)

Lecture. Le poids médian w se situe entre celui de la fréquence (0,049) et celui de la mortalité (0,493) : le contrefactuel repose encore pour l'essentiel (78 à 88 %) sur l'historique propre du site, et le diagnostic du §3.7 (contrefactuel reposant surtout sur l'historique propre du site, retour à la moyenne commun au groupe non corrigé) s'applique donc à la gravité comme à la fréquence. À 500 m, l'Empirical Bayes est plus négatif que le test de cassure de 5 à 6 points,

avec des intervalles largement recouvrants ; à 1 km, les deux estimateurs sont de signes opposés, tous non significatifs. Le KSI à 500 m est la seule de ces quatre estimations dont l'intervalle exclut zéro, alors que la cassure correspondante en est très éloignée : c'est le motif de divergence observé sur les tués, en atténué. Comme sur les tués, l'Empirical Bayes ne peut pas porter ici une conclusion d'existence (§3.7) ; la conclusion « aucun effet détectable » sur la gravité repose sur le test de cassure et sur ses seuils de détection.

Sensibilité au choix de la tendance. Contrairement à la fréquence, où l'indicateur et la tendance coïncident, la mortalité admet deux définitions plausibles de la tendance témoin. Si celle-ci est mesurée sur le nombre d'accidents des zones témoins plutôt que sur leurs tués, l'effet passe de -30,1 % à -21,7 % à 500 m, et de -24,3 % à -11,9 % à 1 km (borne haute de l'IC frôlant zéro). Cette sensibilité de spécification, qui s'ajoute à la divergence EB / test de cassure discutée au §8, illustre que le signal sur la mortalité dépend fortement des choix de modélisation là où les données sont les plus rares ; elle impose la lecture prudente des §8 et §9 (indice non concluant).

Conversion en nombres absolus (rayon 500 m, fenêtre post-installation observée). Les effets relatifs se convertissent en comptes sur la fenêtre observée après installation (2 767 années-zone pour les accidents et les tués, 2017-2024 ; 2 607 pour les blessés graves, 2018-2024). Convention de calcul : pour le test de cassure, le contrefactuel vaut observé / (1 + effet), avec les bornes de l'IC 95 % ; pour l'EB, le bilan vaut $\Pi - \lambda$, avec pour bornes effet $\times \Pi$. La colonne p reprend la valeur de la régression pour le test de cassure (§8) et, pour l'EB, un test de Wald sur θ calculé à partir de son erreur-type (cohérent avec l'IC affiché). Valeurs arrondies (ordres de grandeur).

INDICATEUR (OBSERVÉ APRÈS)	MÉTHODE	P	BILAN CENTRAL	IC 95 %
Accidents (2 679)	Cassure (+3,4 %)	0,52	≈ 90 supplémentaires	190 évités à 340 supplémentaires
Accidents	EB témoin (−3,4 %)	0,20	≈ 90 évités	240 évités à 50 supplémentaires
Accidents	EB trafic (−7,5 %)	0,003	≈ 220 évités	360 à 80 évités
Blessés graves (771)	Cassure (−1,5 %)	0,89	≈ 10 évités	190 évités à 130 supplémentaires
Tués (134)	Cassure (−17 %)	0,38	≈ 27 évités	110 évités à 27 supplémentaires
Tués	EB tendance tués (−30 %)	< 0,001	≈ 57 évités	31 à 84 évités
Tués	EB tendance accidents (−22 %)	0,006	≈ 37 évités	10 à 64 évités

Lecture. Ces fourchettes changent de signe pour les trois indicateurs (selon la méthode, pour les tués) : aucun bilan n'est identifiable à partir de ces données, ni en « vies sauvées » ni en « victimes supplémentaires ». Pour rappel, ces chiffres ne couvrent que l'effet **localisé** (moins de 500 m des radars récents, périmètre où surviennent environ 2 % des tués géolocalisés de la période ; 4,6 % si l'on retient l'ensemble du parc), hors tout effet de réseau (la dissuasion que le contrôle automatisé peut exercer sur l'ensemble du réseau routier, loin des cabines, que l'étude ne mesure pas) ; ils s'entendent en « intention de traiter » (cabine installée, active ou non, §11) ; et ce sont des ordres de grandeur, pas des comptes à l'unité.

13.5 Contrôle d'adoption échelonnée : estimateur de type Callaway-Sant'Anna

L'estimateur à effets fixes bidirectionnels peut être biaisé en adoption échelonnée si les effets varient entre cohortes (annexe 13.3, Limites statistiques). Pour borner ce risque, l'effet a été ré-estimé sur la fréquence des accidents (rayon 500 m) selon la logique de Callaway & Sant'Anna (2021) : un effet $ATT(g, t)$ est estimé séparément pour chaque cohorte g (année d'installation) et chaque année t postérieure, par différence-en-différences de Poisson 2×2 entre l'année de référence $g - 1$ et l'année t (forme fermée : logarithme du rapport croisé des taux d'accidents par année-couverture), puis les $ATT(g, t)$ sont agrégés en pondérant par la taille des cohortes.

L'intérêt de ce découpage est d'écarter une erreur propre aux installations échelonnées. Lorsque les dates de pose diffèrent, une régression classique finit par utiliser des sites **déjà équipés** comme témoins de sites équipés plus tard. Ces sites ne sont plus des témoins valables, puisque leur trajectoire contient déjà l'effet de leur propre radar, et cette comparaison entre dans le résultat d'ensemble avec un poids qui peut en inverser le signe. Ici, chaque $ATT(g, t)$ est calculé contre des zones qui, à la date considérée, n'ont pas encore reçu de radar (variante A) ou n'en ont pas reçu de nouveau (variante B).

L'incertitude est estimée par bootstrap, en rééchantillonnant des **zones entières** plutôt que des couples zone-année : les années d'une même zone partagent la même route et le même trafic, elles ne sont donc pas indépendantes, et les traiter comme telles donnerait un intervalle trop étroit. Le tirage porte sur 999 réplifications à graine fixée, l'intervalle étant donné par les centiles 2,5 et 97,5. Deux groupes de comparaison ont été employés (univers corrigé, §7.3 : 1 002 zones traitées, 1 274 zones témoins).

VARIANTE	GROUPE DE COMPARAISON	CELLULES ATT(G,T)	EFFET AGRÉGÉ	IC 95 % (BOOTSTRAP)
A	zones pas encore traitées	10	-3,2 %	[-15,9 % ; +9,1 %]
B	radars anciens (sans changement ; protocole principal)	15	-0,7 %	[-10,7 % ; +10,4 %]

La variante A perd mécaniquement l'année 2024 (plus aucune zone « pas encore traitée » à cette date). Détail des $ATT(g, t)$ de la variante B, en % :

cohorte 2019	:	2020	-30,7	·	2021	+9,4	·	2022	-5,0	·	2023	-15,2	·	2024	-5,1
cohorte 2020	:	2021	+9,5	·	2022	-0,3	·	2023	+8,2	·	2024	+6,2			
cohorte 2021	:	2022	+2,2	·	2023	+1,3	·	2024	+6,0						
cohorte 2022	:	2023	-11,7	·	2024	-3,7									
cohorte 2023	:	2024	-3,0												

Lecture. Aucun effet significatif : les deux variantes donnent un effet agrégé proche de zéro, cohérent avec le test de cassure (+3,4 %) et l'Empirical Bayes (-3,4 %), pour un effet minimal détectable d'environ 14 à 17 % à la baisse, comparable à celui du test principal (ce seuil se calcule sur l'échelle logarithmique du modèle, comme à l'annexe 13.3 : erreur-type = $[\ln(1 + \text{borne haute}) - \ln(1 + \text{borne basse})] / (2 \times 1,96)$ à partir de l'intervalle bootstrap, puis seuil = $1 - \exp(-2,80 \times \text{erreur-type})$, soit 17,0 % pour la variante A et 14,0 % pour la variante B). Deux précisions d'interprétation. D'abord, cet estimateur ne corrige pas la régression vers la moyenne : l'année de référence $g - 1$ appartient à la période du pic de sélection, ce qui oriente les ATT vers le négatif ; le léger penchant négatif observé (-1 à -3 %) est compatible avec ce reflux résiduel. Ensuite, et c'est l'objet du contrôle : même sans correction de pente, un estimateur robuste à l'hétérogénéité des cohortes ne révèle aucun effet que le TWFE aurait masqué par ses « comparaisons interdites ». La conclusion du §8 sur la fréquence (effet localisé non identifié) n'est donc pas imputable au biais d'adoption échelonnée. Ce contrôle ne se soustrait pas pour autant à la réserve du §8 : appliqué à la variante B, le test placebo ne restitue pas davantage zéro (écart médian de 7,4 points sur quatre configurations, une significative ; annexes 13.6, section E, et 13.7).

13.6 Sensibilité de l'estimation principale

Cette annexe publie les estimations obtenues sur les mêmes données et le même échantillon, en faisant varier les choix de spécification qui ne sont pas déterminés par les données. Sa conclusion est que l'estimation ponctuelle de l'effet localisé n'est pas invariante à ces choix ; le lecteur doit en tenir compte avant de retenir un chiffre isolé.

Degrés de liberté du chercheur. Cette étude n'a pas fait l'objet d'un préenregistrement, ce qui est assumé et doit être pris en compte dans la lecture des variantes qui suivent. La spécification principale (régression de Poisson à tendance linéaire et cassure à l'installation, rayon 500 m, exclusion de l'année d'installation, erreurs-types regroupées par zone) a été arrêtée avant la conduite du test placebo de l'annexe 13.7. C'est ce test, postérieur, qui fonde la conclusion de non-identification sur la fréquence : elle ne résulte pas d'un choix de spécification opéré après examen des résultats. Les annexes 13.4 à 13.7 publient plus d'une centaine d'estimations, y compris celles qui déplacent l'estimation ponctuelle de plusieurs dizaines de points et celles qui en inversent le signe ; aucune n'a été écartée en raison de son résultat. Nos calculs en produisent davantage : parmi les sorties non reprises ici figurent l'Empirical Bayes à poids forcé $w = 1$ (le cas $w = 0$ figure à l'annexe 13.4), le détail des cellules ATT(g, t) de la variante A et les tests d'écart entre sous-groupes de l'étape 17 autres que ceux cités à l'annexe 13.6, section E ; les découpages par cohorte des configurations placebo, sous-échantillons emboîtés de la première configuration, figurent à l'annexe 13.7 (décomposition par cohorte), et celui restreint aux cohortes 2023-2024 donne +7,9 % ($p = 0,505$) pour la spécification retenue. Toutes sont produites par les étapes 8, 10, 12 et 17 du programme de reproduction, publié avec ses sorties, qui permet de reproduire chacune d'elles et d'en estimer d'autres.

A. Forme imposée au reflux. Le modèle de l'annexe 13.3 écrit la trajectoire de reflux comme linéaire en temps relatif, sur tout le support. La théorie du retour à la moyenne prédit en réalité un reflux amorti (convexe), et rien dans les données ne fixe la forme. Accidents, 500 m, 2017-2024, même échantillon :

SPÉCIFICATION DE LA TENDANCE	EFFET	P	IC 95 %
Linéaire sur tout le support (retenue au §8)	+3,4 %	0,52	[-6,7 ; +14,7]
Quadratique	+19,6 %	0,022	[+2,6 ; +39,5]
Cubique	+31,3 %	0,004	[+8,9 ; +58,2]
Pente estimée sur les seules années antérieures	-16,1 %	0,002	[-25,0 ; -6,2]
Sans pente (DiD naïf, cf. annexe 13.1)	-28,9 %	< 0,001	[-39,5 ; -16,3]
Linéaire, hors année 2020	+13,4 %	0,044	[+0,4 ; +28,1]
Linéaire, fenêtre événementielle [-2 ; +2]	-14,4 %	0,155	[-31 ; +6]
Linéaire, fenêtre événementielle [-3 ; +3]	+4,1 %	0,616	[-11 ; +22]

À 1 km, la structure est la même (+4,9 % linéaire, +21,2 % quadratique, -17,7 % pente antérieure, -33,4 % sans pente). Sous la variante « pente antérieure seule », les autres indicateurs à 500 m donnent : tués -23,1 % (p = 0,11), blessés graves -13,1 % (p = 0,10), KSI -13,8 % (p = 0,065) ; à 1 km, -17,2 %, -9,1 % et -9,9 %.

Lecture. Imposer une pente constante prolonge après la pose un reflux qui devrait s'être amorti : le contrefactuel est alors trop bas et l'effet estimé tiré vers le positif. Interrompre le reflux à la date de pose fait l'inverse. Les données ne tranchent pas entre les deux. La ligne « hors année 2020 » relève d'un autre mécanisme : retirer l'année du creux de circulation fait passer l'estimation de +3,4 % (p = 0,52) à +13,4 % (p = 0,044), et les configurations placebo excluant 2020 produisent les écarts à zéro les plus grands de la spécification retenue ; le choc de circulation de 2020-2021 est discuté comme limite au §11. Ce point est décisif pour la lecture du §3.7 : sous la variante « pente antérieure seule », les ordres de grandeur obtenus (13 à 23 % à 500 m, 9 à 18 % à 1 km) recouvrent en partie ceux de la littérature corrigée (12 à 22 %, §3.5). Cette variante n'a pas pour autant le statut d'une estimation rivale de même valeur : 420 des 1 002 zones traitées ne disposent que de deux ou trois années antérieures à la pose, et elle ne restitue pas davantage zéro au test placebo (écart médian de 9,3 points, négatif dans les sept configurations, annexe 13.7). Le seuil détectable sur la fréquence à 500 m varie également avec la spécification : 13,7 % à la baisse et 15,9 % à la hausse (linéaire), 19,7 % et 24,5 % (quadratique), 14,8 % et 17,3 % (pente antérieure seule).

B. Étude d'événement non contrainte. L'annexe 13.1 regroupe les temps relatifs en quatre blocs. Le même modèle avec des indicatrices année par année (référence : -1 an) donne :

TEMPS RELATIF	-7	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5
Coefficient	+0,02	+0,14	+0,38	+0,27	+0,02	+0,02	réf.	-0,12	-0,08	-0,50	-0,64	-0,36
p	0,92	0,34	<0,001	<0,001	0,78	0,72	réf.	0,018	0,29	<0,001	<0,001	0,078
zones	111	265	414	582	871	1 002	1 002	891	737	588	420	131

La trajectoire n'est pas continûment descendante : elle est élevée à -5 et -4 ans, revenue au niveau de référence à -3, -2 et -1, stable à +1 et +2, puis décroche à +3 et +4. Les effectifs de la dernière ligne signalent la raison de cette forme : les temps relatifs éloignés ne sont renseignés que par une fraction des zones (111 sur 1 002 à -7 ans), et cette fraction est déterminée par la cohorte d'installation. Les temps -5 et -4 sont peuplés de cohortes tardives observées en 2017-2018, les temps +3 et +4 de cohortes 2019-2020 observées en 2022-2024 : les coefficients mélangent donc temps relatif et composition.

Ce point est tranché par le placebo de l'annexe 13.7, qui estime le profil pré-installation à **composition constante sur les cinq années qui précèdent la pose**, sur les seules cohortes 2022-2024. Ce profil est monotone **croissant** et culmine l'année précédant la pose ; le coefficient à -5 ans y vaut -0,14, contre +0,39 dans la même coupe poolée restreinte aux années antérieures à la pose, sur exactement les mêmes zones (le +0,38 du tableau ci-dessus provient de l'étude d'événement sur le panneau complet, années postérieures comprises : les deux modèles diffèrent par le support, non par l'échantillon). La trajectoire réelle monte donc jusqu'à l'installation puis redescend, ce qui est bien la signature d'une sélection réactive sur un pic, mais non celle d'un reflux déjà engagé plusieurs années auparavant, comme la lecture poolée le laissait croire. C'est aussi la raison mécanique pour laquelle aucune tendance monotone ne peut épouser correctement cette trajectoire, et pourquoi le signe de l'effet estimé dépend de la forme retenue.

C. Autres variantes de spécification (accidents, tendance linéaire, rayon de 500 m sauf mention contraire) :

VARIANTE	EFFET	P	IC 95 %	SEUIL DÉTECTABLE (BAISSE / HAUSSE)	N
Spécification retenue au §8	+3,4 %	0,52	[-6,7 ; +14,7]	13,7 % / 15,9 %	17 206
Univers complet des zones, comptages nuls inclus (1 521 traitées, 1 646 témoins)	+6,6 %	0,23	[-3,9 ; +18,2]	13,7 % / 15,9 %	23 815
Binomiale négative au lieu de Poisson ($\alpha = 3,54$), erreurs-types cluster	+6,4 %	0,22	[-3,7 ; +17,5]	13,3 % / 15,3 %	17 206
Traités restreints aux tourelles, témoins aux cabines fixes et discriminants	+0,6 %	0,91	[-10,0 ; +12,6]	14,8 % / 17,3 %	11 429
Exclusion des témoins mis en service en 2017 ou 2018 (36 des 1 274 zones témoins de l'estimation, soit 52 des 1 646 radars du fichier source)	+3,6 %	0,50	[-6,5 ; +14,9]	13,7 % / 15,9 %	16 918
Témoins restreints aux radars mis en service avant 2013 (923 des 1 274 zones témoins)	+3,9 %	0,47	[-6,3 ; +15,3]	13,8 % / 16,0 %	14 398
Témoins restreints aux radars mis en service avant 2010 (343 des 1 274 zones témoins)	+3,8 %	0,48	[-6,4 ; +15,1]	13,8 % / 15,9 %	9 758
Exclusion des témoins ayant reçu un radar neuf à moins de 500 m (37 des 1 274 zones témoins)	+3,2 %	0,55	[-6,9 ; +14,4]	13,7 % / 15,9 %	16 910
Exclusion des témoins ayant reçu un radar neuf à moins de 1 km (109 des 1 274 zones témoins)	+2,8 %	0,61	[-7,3 ; +13,9]	13,7 % / 15,9 %	16 334
Erreurs-types regroupées par maille spatiale ($\approx 1^\circ$ de latitude/longitude) au lieu de la zone	+3,4 %	0,58	[-8,2 ; +16,6]	15,7 % / 18,7 %	17 206
Rayon de 250 m au lieu de 500 m (736 zones traitées, 1 041 témoins)	-7,1 %	0,35	[-20,3 ; +8,4]	19,8 % / 24,7 %	13 480

Quatre remarques. (i) Le filtre « au moins un accident géolocalisé sur la fenêtre » (annexe 13.4) est une sélection sur la variable expliquée, appliquée de façon asymétrique (66 % des traités, 77 % des témoins) ; la première ligne du tableau en mesure l'effet. (ii) La surdispersion étant forte, la famille de comptage déplace l'estimation ponctuelle de 3 points, sans changer la conclusion. (iii) Le regroupement spatial des erreurs-types les augmente d'environ 15 % : les seuils détectables publiés au §8 et au §9 sont donc de légers minorants : sur la fréquence à 500 m, 13,7 % passe à 15,7 % à la baisse et 15,9 % à 18,7 % à la hausse. (iv) Les quatre variantes portant sur le groupe témoin (ancienneté, exposition à un radar neuf) déplacent l'estimation de moins d'un point et les seuils d'au plus un dixième de point : l'information vient des zones traitées, et ces contrôles écartent qu'un reflux résiduel des témoins récents ou leur exposition partielle porte le résultat, sans rien dire de plus sur la spécification. Le rayon de 250 m est une sensibilité, non un test : le seuil détectable y passe de 13,7 % à 19,8 % à la baisse, l'intervalle contient l'estimation ponctuelle de la spécification retenue et recouvre largement le sien (de -6,7 % à +8,4 %), et le changement de signe du point estimé (10,5 points d'écart) reste inférieur à ce seuil.

D. Composition technologique des groupes. La composition est disjointe. Groupe traité (1 521 radars installés 2019-2024) : 1 328 tourelles (87 %), 178 radars discriminants, 13 radars urbains, 2 cabines fixes classiques. Groupe témoin (1 646 radars antérieurs à 2019) : 637 cabines fixes, 529 radars de feu rouge (32 %), 339 discriminants, 71 passages à niveau, 70 radars tronçon. L'étude porte donc en pratique, à 87 %, sur les radars « tourelles », et le groupe témoin comprend des dispositifs dont la zone d'influence se prête mal à un disque de 500 m (tronçon, passage à niveau). L'hypothèse de tendances parallèles est ainsi formulée entre deux populations de sites qui diffèrent par la technologie autant que par l'ancienneté. Le contrôle en groupes homogènes (tableau C ci-dessus, +0,6 %, $p = 0,91$) ne fait pas apparaître d'effet de composition, dans la limite de ce qu'il peut distinguer : sur 11 429 lignes, son intervalle de confiance va de -10,0 % à +12,6 % et son seuil de détection vaut 14,8 % à la baisse et 17,3 % à la hausse.

Comparabilité des groupes sur d'autres dimensions que la technologie. Le tableau ci-dessous compare les deux groupes sur trois dimensions non technologiques : la vitesse maximale autorisée au radar (fichier des radars), caractéristique de la **zone** ; la part des accidents survenus en agglomération (champ agg des fichiers caractéristiques) et leur répartition par catégorie de route (champ catr des fichiers lieux), qui décrivent l'une et l'autre les **accidents** observés dans la zone sur 2017-2024, avant et après pose confondus, et non la zone elle-même. Les unités d'observation diffèrent donc d'un bloc à l'autre et les blocs ne se lisent pas sur le même plan. La VMA n'est par ailleurs pas renseignée pour tous les radars : elle manque pour 97 des 1 002 zones traitées et pour 544 des 1 274 zones témoins, ces dernières étant très majoritairement des radars de feu rouge (491) et des passages à niveau (51), dispositifs auxquels le fichier n'associe aucune vitesse limite. Les pourcentages du premier bloc sont donc calculés sur les seules zones renseignées, et ne décrivent qu'un peu plus de la moitié du groupe témoin.

DIMENSION	MODALITÉ	ZONES TRAITÉES	ZONES TÉMOINS
VMA du radar (unité : la zone)	<i>zones à VMA renseignée</i>	<i>905 sur 1 002 (90,3 %)</i>	<i>730 sur 1 274 (57,3 %)</i>
	30 km/h	0 (0,0 %)	6 (0,8 %)
	50 km/h	213 (23,5 %)	106 (14,5 %)
	60 km/h	0 (0,0 %)	1 (0,1 %)
	70 km/h	149 (16,5 %)	135 (18,5 %)
	80 km/h	265 (29,3 %)	107 (14,7 %)
	90 km/h	205 (22,7 %)	170 (23,3 %)
	110 km/h	59 (6,5 %)	141 (19,3 %)
	130 km/h	14 (1,5 %)	64 (8,8 %)
Milieu (unité : l'accident)	<i>accidents à moins de 500 m</i>	<i>9 761</i>	<i>35 054</i>
	hors agglomération	3 261 (33,4 %)	6 180 (17,6 %)
	en agglomération	6 500 (66,6 %)	28 874 (82,4 %)
Catégorie de route (unité : l'accident)	<i>accidents à moins de 500 m</i>	<i>9 761</i>	<i>35 054</i>
	autoroute	1 621 (16,6 %)	4 188 (11,9 %)
	route nationale	822 (8,4 %)	1 691 (4,8 %)
	route départementale	2 854 (29,2 %)	6 229 (17,8 %)
	voie communale	3 781 (38,7 %)	21 806 (62,2 %)
	route de métropole urbaine	549 (5,6 %)	797 (2,3 %)
	autre ou non renseignée	134 (1,4 %)	343 (1,0 %)

Lecture. Sur la VMA renseignée, les zones traitées se concentrent entre 50 et 90 km/h, alors que les zones témoins renseignées comptent une part plus élevée de sites à 110 et 130 km/h (19,3 % et 8,8 %, contre 6,5 % et 1,5 %). Sur les accidents, la part survenue en agglomération est de 66,6 % dans les zones traitées et de 82,4 % dans les zones témoins, et la part survenue sur voie communale de 38,7 % contre 62,2 %. Ces deux dernières dimensions situent le groupe témoin plus près du milieu urbain, ce que sa composition technologique laissait attendre ; la première, qui ne porte que sur les zones renseignées, le situe davantage sur le réseau rapide. Les deux constats ne s'opposent pas : ils reposent sur des unités différentes, la zone d'un côté et l'accident de l'autre, et sur des sous-populations différentes. Au niveau du

fichier des radars (1 521 traités, 1 646 témoins), les ordres de grandeur sont les mêmes : VMA renseignée pour 93,3 % des traités et 63,4 % des témoins, dont 5,4 % et 2,2 % à 110 et 130 km/h côté traité, contre 19,0 % et 9,5 % côté témoin.

Il s'agit d'une description, non d'une stratification. Aucun de ces découpages ne sert à estimer, corriger ou pondérer quoi que ce soit : les estimations du §8 et leurs intervalles sont inchangés. L'argument de puissance exposé au §11 continue de justifier l'absence de stratification estimative sur ces mêmes dimensions. Ces chiffres sont produits par l'étape 13 du programme de reproduction.

E. Nature du traitement et diagnostic de spécification par sous-groupe. Le §5 établit, par rapprochement du fichier des radars avec une extraction du même fichier antérieure aux dégradations de l'hiver 2018-2019, que 1 030 des 1 521 radars datés 2019-2024 (67,7 %) remplacent un radar présent à moins de 100 m en octobre 2018 et absent du fichier actuel, que 19 (1,2 %) cohabitent avec un radar de 2018 toujours listé, et que 472 (31,0 %) sont des sites nouveaux (étape 16 du programme de reproduction). Au seuil de 25 m la part de rééquipements stricts vaut 52,1 %, à 50 m 59,6 %, à 150 m 71,1 % et à 250 m 72,6 % : au-delà de 50 m la part varie peu (60 à 73 %), et le seuil de 25 m est de l'ordre de la précision des coordonnées. Le fait qu'une majorité de radars récents remplacent un radar existant ne dépend donc pas du seuil retenu, la valeur exacte de la part en dépend. Sur les 1 002 zones traitées de l'estimation à 500 m, 658 sont des sites rééquipés, 17 des cohabitations et 327 des sites nouveaux ; à 1 km, 869, 18 et 403. Cette variable ne sert à aucune estimation publiée. Elle a été utilisée pour un diagnostic, dont il faut d'abord dire pourquoi il ne produit pas de résultat publiable, puis ce qu'il montre (étape 17 du programme de reproduction ; les cohabitations, ni nouvelles ni rééquipées, en sont écartées et ne figurent que dans l'ensemble).

Ce que donne la spécification retenue au §8, appliquée séparément aux deux sous-groupes. Accidents, 2017-2024, témoins communs (1 274 zones à 500 m, 1 470 à 1 km) :

SOUS-GROUPE, RAYON	ZONES	CASSURE θ (P, IC 95 %)	SEUILS BAISSSE / HAUSSE	PENTE EVT_LIN	DID SANS PENTE	DID AVEC EFFETS FIXES DE COHORTE
Ensemble, 500 m	1 002	+3,4 % (0,52 ; [-6,7 ; +14,7])	13,7 / 15,9 %	-0,074	-28,9 % (< 0,001)	-0,7 % (0,86 ; [-8,2 ; +7,4])
Sites nouveaux, 500 m	327	-18,7 % (0,005 ; [-29,6 ; -6,2])	18,6 / 22,8 %	+0,042	+0,6 % (0,97)	-1,7 % (0,79 ; [-13,3 ; +11,5])
Sites rééquipés, 500 m	658	+28,0 % (< 0,001 ; [+10,6 ; +48,2])	18,9 / 23,3 %	-0,128	-33,3 % (< 0,001)	+2,0 % (0,69 ; [-7,4 ; +12,4])
Ensemble, 1 km	1 290	+4,9 % (0,27 ; [-3,6 ; +14,1])	11,3 / 12,8 %	-0,089	-33,4 % (< 0,001)	-1,1 % (0,71 ; [-6,9 ; +5,0])
Sites nouveaux, 1 km	403	-14,4 % (0,015 ; [-24,5 ; -2,9])	16,5 / 19,7 %	+0,025	-2,9 % (0,82)	+0,6 % (0,90 ; [-8,3 ; +10,4])
Sites rééquipés, 1 km	869	+26,1 % (< 0,001 ; [+12,2 ; +41,6])	15,3 / 18,1 %	-0,138	-37,5 % (< 0,001)	-0,5 % (0,89 ; [-7,7 ; +7,1])

Lus isolément, -18,7 % et +28,0 % se présenteraient comme deux effets significatifs de sens opposé. Ils ne le sont pas, pour quatre raisons que le programme imprime.

1. *L'écart vient du terme de pente, non de deux trajectoires.* L'identité $\theta_{\text{cassure}} = \theta_{\text{DiD}} - \text{pente} \times$ (décalage moyen d'événement entre années postérieures et antérieures) se vérifie dans les six lignes : pour les sites nouveaux à 500 m, la part DiD vaut +0,006 en logarithme, le produit pente $\times$ décalage ($0,042 \times 5,10$) vaut +0,214 et se retranche, d'où un θ prédit de -0,211 sur les valeurs non arrondies, contre -0,208 observé (soit le -18,7 %) ; pour les rééquipés, +0,251 prédit contre +0,247 observé. Or, faute d'effet fixe de zone, et evt valant année moins cohorte avec des indicatrices d'année saturées, la pente evt_lin n'est pas identifiée par la trajectoire d'un site : elle capte l'inverse du gradient de niveau entre cohortes de pose. Les rééquipés tardifs (2021-2024) occupent des sites plus accidentogènes que ceux de 2019-2020, d'où une pente négative extrapolée en hausse contrefactuelle, et inversement pour les nouveaux.
2. *Avec des effets fixes de cohorte, le modèle sans pente converge vers zéro* (dernière colonne : -0,7 %, -1,7 %, +2,0 % à 500 m). Ajoutés à la spécification publiée, pente comprise, ces effets fixes ne suffisent pas : l'écart entre sous-groupes y reste de -25,5 % ($p = 0,005$) à 500 m et de -16,3 % ($p = 0,059$) à 1 km, sur le modèle empilant les deux sous-groupes.
3. *Les estimateurs qui n'ont pas cette pente ne distinguent pas les sous-groupes.* Empirical Bayes (tendance témoin mesurée sur les accidents, annexe 13.4) : sites nouveaux -2,5 % [-11,0 ; +6,0] (269 sites, poids médian 0,035), rééquipés -2,6 % [-9,3 ; +4,1] (605 sites, poids 0,070), écart +0,1 point ($p = 0,99$). Callaway-Sant'Anna, variante B : -2,6 % [-21,1 ; +21,5] contre -1,3 % [-13,3 ; +13,1], écart -1,4 % (p bootstrap 0,81). Le DiD sans pente les sépare, lui, dans l'autre sens : +47,5 % en rapport des taux ($p = 0,022$), écart estimé sur

un modèle empilant les deux sous-groupes, dont les points diffèrent légèrement de ceux du tableau ci-dessus, estimés séparément. Le contraste dépend donc entièrement de l'estimateur.

4. *Le placebo échoue sur les sites nouveaux plus nettement que sur l'ensemble.* Spécification linéaire, sept configurations de l'annexe 13.7 (fausse date g-2 ou g-3, cohortes 2022-2024, effet nul par construction) :

CONFIGURATION	ENSEMBLE	SITES NOUVEAUX	SITES RÉÉQUIPÉS
g-2, 500 m	-11,2 % (0,25 ; 414 zones)	-30,4 % (0,004 ; 220)	+11,3 % (0,45 ; 186)
g-2, 1 km	-19,1 % (0,034 ; 507)	-35,3 % (< 0,001 ; 271)	+2,4 % (0,87 ; 227)
g-3, 500 m	-19,2 % (0,14)	-49,7 % (< 0,001)	+25,1 % (0,18)
g-3, 1 km	-20,5 % (0,081)	-39,0 % (0,006)	+1,0 % (0,96)
g-2, 500 m, univers restreint	-13,9 % (0,14 ; 374)	-31,5 % (0,002 ; 200)	+5,8 % (0,69 ; 167)
g-2, 500 m, hors 2020	-30,8 % (0,012)	-51,7 % (< 0,001)	+1,6 % (0,94)
g-3, 500 m, hors 2020	-44,4 % (0,038)	-74,9 % (< 0,001)	+17,8 % (0,61)
Écart absolu médian, significatifs, signe	19,2 pts, 3/7, 7 négatifs	39,0 pts, 7/7, 7 négatifs	5,8 pts, 0/7, 7 positifs

Sur les sites nouveaux, dans une fenêtre où aucun radar n'est posé, la spécification renvoie des écarts de même signe et d'amplitude supérieure à l'estimation réelle du sous-groupe ; le -18,7 % est donc, plus encore que le +3,4 % de l'ensemble, une valeur que le seul biais de spécification suffit à produire. Sur les rééquipés, le placebo est proche de zéro, mais avec des seuils de détection propres de 32 à 60 % à la baisse et de 48 à 149 % à la hausse selon la configuration, supérieurs au +28,0 % à expliquer : il classe la spécification, il ne la certifie pas. Le placebo couvre 67 % des zones qui portent le -18,7 % et 28 % de celles qui portent le +28,0 % ; son écart ne se retranche pas des estimations, les échantillons et les supports différant (annexe 13.7).

Gravité et mortalité. Le même dessin se retrouve, avec des effectifs plus faibles encore. Sites nouveaux, 500 m : tués 2017-2024 -52,7 % (p = 0,014 ; 148 → 27 tués sur les 327 zones ; seuil détectable 57,4 % à la baisse, que l'estimation n'atteint pas), -44,5 % (p = 0,115) sur 2018-2024 ; blessés graves -34,7 % (p = 0,010 ; 621 → 181) ; KSI -37,0 % (p = 0,003) ; à 1 km -29,0 %, -24,2 % et -23,7 %. Sites rééquipés : tués +12,1 % (p = 0,69 ; 148 → 107), blessés graves +23,6 % (p = 0,13 ; 737 → 568) à 500 m et +31,1 % (p = 0,003 ; 1 835 → 1 442) à 1 km, KSI +23,8 % et +27,4 %. Le placebo sur ces issues (douze configurations) est significatif dix fois sur douze chez les sites nouveaux (écart médian 45 points, toujours négatif) et zéro fois

chez les rééquipés (8,2 points). Un fait descriptif, lui, ne dépend d'aucune spécification : l'écart entre sous-groupes vient de la gravité conditionnelle et non du nombre d'accidents ; les tués pour 100 accidents passent de 8,53 en 2017 à 3,83 en 2024 dans les zones des sites nouveaux, par une série annuelle non monotone dont le minimum est 1,08 en 2023, calculée sur 422 à 610 accidents par an, de 4,62 à 6,04 dans celles des sites rééquipés, de 2,10 à 2,07 chez les témoins et de 7,60 à 7,52 pour les accidents survenus à plus de 1 km de tout radar. Les sites nouveaux partent d'un niveau de gravité supérieur à tous les groupes de référence et convergent, ce qui est compatible avec une sélection des sites sur la gravité récente, hypothèse que les BAAC, qui ne contiennent pas le motif d'implantation, ne permettent pas de tester.

Deux faits annexes sortis du même diagnostic. Le placebo appliqué à l'Empirical Bayes (quatre configurations, g-2 et g-3 aux deux rayons, cohortes 2022-2024) donne sur l'ensemble +12,5 %, +7,2 %, +12,8 % et +10,0 %, tous significatifs, écart médian 11,2 points ; appliqué à Callaway-Sant'Anna B, +14,1 %, +4,3 %, +9,3 % et +5,5 %, un significatif, écart médian 7,4 points : dans une fenêtre antérieure à la pose réelle, l'un et l'autre trouvent l'accidentalité de l'année qui précède la pose supérieure à ce que l'historique du site ou les zones à radar ancien laissent attendre, ce qui est le profil de sélection réactive de l'annexe 13.7 et confirme le statut de borne basse et de diagnostic donné à l'Empirical Bayes au §3.7. Enfin, le terme d'exposition de couverture ne déplace pas la cassure d'ensemble ($\theta = +0,033788$ avec, $+0,033744$ sans, écart de 4×10^{-5}), ce qui documente sa redondance avec les indicatrices d'année (§7.3).

Ce que ce diagnostic ne dit pas. Il n'établit ni que l'effet diffère entre sites nouveaux et sites rééquipés, ni qu'il ne diffère pas ; il établit que la spécification du §8 ne permet pas d'en décider, ce qui est la conclusion du §8 elle-même, étendue aux deux sous-groupes. Sur la gravité et la mortalité, la conclusion publiée au §10, aucun effet détectable, reste inchangée et vaut pour les deux sous-groupes : les seuils de détection y sont plus larges encore que sur l'ensemble, et le placebo échoue dix fois sur douze sur les sites nouveaux. Le partage n'était pas prévu au protocole, les p-valeurs de cette section ne sont corrigées d'aucune multiplicité, et le statut du site est confondu avec le millésime de pose (à 500 m, 90 % des zones de la cohorte 2019 et 92 % de celles de 2020 portaient un radar en 2018, 58 % en 2021, 45 à 48 % de 2022 à 2024), donc avec la durée de suivi et le niveau d'accidentalité des cohortes.

13.7 Test placebo temporel

Ce test répond à une question simple : le dispositif d'estimation produit-il zéro là où l'effet vaut zéro ? C'est lui qui fonde la réserve énoncée au §8 sur la fréquence des accidents.

Protocole. On attribue à chaque radar une **fausse date de mise en service**, antérieure de deux ans (puis de trois ans) à sa date réelle, et on restreint l'échantillon aux années **strictement antérieures à l'installation réelle**. Dans cette fenêtre, aucun des radars étudiés n'a encore été posé (le site pouvant, pour les rééquipés, porter un radar plus ancien, §5) : l'effet à estimer vaut zéro par construction. Le test porte sur les cohortes 2022-2024, seules à disposer d'une antériorité suffisante (5 à 7 années avant traitement). L'année de la fausse pose est exclue, comme l'année d'installation l'est dans l'analyse principale ; les témoins, l'offset de couverture GPS et l'inférence par erreurs-types regroupées par zone sont ceux du modèle principal.

Résultats : 0 placebo, fréquence des accidents (valeur idéale : 0,0 %). Les sept configurations du tableau sont celles qui ne sont pas des découpages par cohorte de la première ; trois d'entre elles (univers restreint, hors 2020 à g-2 et à g-3) restent emboîtées dans les quatre premières, de sorte que les sept ne constituent pas sept épreuves indépendantes. Nos calculs en produisent douze : quatre sont des découpages par cohorte de la première configuration (cohortes 2023-2024, dont le résultat est cité à l'annexe 13.6, puis 2022, 2023 et 2024 séparément), repris plus bas sous « Décomposition par cohorte » ; une porte sur les tués et figure sous « Sur la mortalité, le placebo est sans portée ». L'écart médian de la synthèse porte sur ces sept configurations, les découpages par cohorte n'y étant pas comptés une seconde fois.

CONFIGURATION	LINÉAIRE (RETENUE AU §8)	QUADRATIQUE	PENTE ANTÉRIEURE SEULE	SANS PENTE (DiD)
g-2, 500 m (414 zones, 10 950 lignes)	-11,2 % (0,248)	+49,9 % (0,086)	-3,2 % (0,636)	+19,1 % (0,004)
g-2, 1 km (507 zones, 12 790 lignes)	-19,1 % (0,034)	+10,2 % (0,619)	-10,6 % (0,083)	+14,0 % (0,013)
g-3, 500 m	-19,2 % (0,142)	-7,5 % (0,567)	-7,5 % (0,381)	+20,4 % (0,003)
g-3, 1 km	-20,5 % (0,081)	-10,0 % (0,291)	-9,3 % (0,225)	+17,5 % (0,007)
g-2, 500 m, univers restreint	-13,9 % (0,142)	+51,7 % (0,078)	-5,0 % (0,456)	+20,4 % (0,002)
g-2, 500 m, hors 2020	-30,8 % (0,012)	-9,6 % (0,780)	-20,3 % (0,025)	+14,1 % (0,017)
g-3, 500 m, hors 2020	-44,4 % (0,038)	-50,0 % (0,061)	-22,6 % (0,090)	+8,6 % (0,107)

SYNTHÈSE SUR LES SEPT CONFIGURATIONS

SPÉCIFICATION	ÉCART ABSOLU MÉDIAN À ZÉRO	ÉCARTS SIGNIFICATIFS À 5 %	SIGNE
Pente antérieure seule	9,3 pts	1 / 7	7 négatifs
Quadratique	10,2 pts	0 / 7	3 positifs, 4 négatifs
Sans pente (DiD)	17,5 pts	6 / 7	7 positifs
Linéaire (retenue au §8)	19,2 pts	3 / 7	7 négatifs

Lecture. La spécification retenue au §8 s'écarte de zéro de 19,2 points en médiane, avec un signe constant sur les sept configurations, et significativement dans trois d'entre elles. Cet écart médian excède le seuil de détection annoncé pour cette spécification (13,7 %) et vaut 1,8 fois la demi-largeur de son intervalle de confiance publié (10,7 points, pour un intervalle [-6,7 ; +14,7] large de 21,4 points). Elle n'est donc pas non biaisée. Les autres spécifications à pente ne le sont pas davantage : la variante à pente antérieure produit un écart médian de 9,3 points, négatif dans les sept configurations ; la quadratique est erratique (± 50 %).

Classement, et pourquoi la spécification est maintenue. Sur le critère de ce tableau, l'écart absolu médian à zéro, la spécification retenue au §8 arrive dernière des quatre (19,2 points, contre 17,5 pour le DiD sans pente, 10,2 pour la quadratique et 9,3 pour la pente antérieure seule). Elle est néanmoins conservée comme estimation principale, pour trois raisons qu'il vaut mieux énoncer que laisser deviner. D'abord, ce classement ne certifie aucune spécification (limite 1 ci-dessous) : les seuils de détection du placebo lui-même sont de l'ordre de 25 %

dans les configurations principales, et plus élevés encore dans les variantes restreintes, de sorte qu'un écart médian de 9,3 points ne s'y distingue pas d'un écart de 19,2 points. Ensuite, le signe de l'écart de la spécification linéaire est ici mécaniquement prévisible : la fenêtre placebo ne contient que la branche montante du profil antérieur à la pose (paragraphe Diagnostic), quand la fenêtre principale contient la montée et le reflux. Enfin, la variante la mieux classée, la pente estimée sur les seules années antérieures, est sous-identifiée sur l'échantillon principal, où 420 zones sur 1 002 (cohortes 2019 et 2020) ne disposent que de deux ou trois années antérieures pour estimer cette pente : son meilleur comportement au placebo, obtenu sur les cohortes 2022-2024 qui disposent de cinq à sept années, ne se transpose pas. Le maintien de la spécification linéaire ne vaut donc pas certificat : c'est parce qu'aucune spécification n'est validée que la conclusion du §8 sur la fréquence est un effet localisé non identifié, et non une mesure d'effet.

Décomposition par cohorte (g-2, 500 m, spécification retenue) : +4,9 % pour la cohorte 2022, +34,4 % pour 2023, -13,9 % pour 2024. Aucune cohorte ne porte le résultat poolé et il n'y a pas de valeur aberrante à écarter : l'estimation est simplement instable d'une cohorte à l'autre.

Un acquis, en sens inverse. La comparaison sans correction de tendance (DiD) produit un écart à zéro compris entre +8,6 % et +20,4 %, significatif dans six configurations sur sept (de +14,0 % à +20,4 %), dans une fenêtre antérieure à toute pose réelle ; la septième, g-3 à 500 m hors 2020, donne +8,6 % ($p = 0,107$). C'est la démonstration directe, sans aucune hypothèse sur la forme de la tendance qui précède la pose, que l'accidentalité d'un site est significativement plus élevée l'année précédant la pose que les années antérieures du même site. Cet acquis repose toutefois sur les effets fixes d'année pour neutraliser le creux de circulation de 2020 et le rebond de 2021, dont le §11 rappelle qu'ils n'ont pas été proportionnels selon le réseau : lorsque l'année 2020 est retirée, l'écart passe de +20,4 % à +8,6 % ($p = 0,107$) à g-3 et de +19,1 % à +14,1 % ($p = 0,017$) à g-2. L'essentiel du -30 % du §6 n'est donc pas attribuable au radar, et ce résultat-là ne dépend d'aucun choix de spécification de la tendance ; sa décomposition entre sélection réactive, changement de pondération et tendance du groupe témoin est discutée au §3.7 et en annexe 13.4.

LIMITES PROPRES AU PLACEBO

1. Le placebo **classe** les spécifications, il n'en **certifie** aucune : ses propres seuils de détection sont du même ordre de grandeur que les écarts qu'il mesure. Ils ne sont repris ici que pour les deux configurations principales : pour la spécification retenue au §8, 25,1 % à la baisse et 33,5 % à la hausse à 500 m, 24,4 % et 32,2 % à 1 km ; pour la

quadratique, 48,3 % et 93,5 % à 500 m ; pour la pente antérieure seule, 17,7 % et 21,6 %. Ils sont plus élevés encore dans les variantes à effectif réduit. Un test incapable de voir un biais de 25 % ne peut valider une estimation de +3,4 %. Les seuils de chaque configuration et de chaque spécification, aux deux bornes, sont produits par l'étape 10 du programme de reproduction et reportés dans les sorties publiées avec lui (section « Disponibilité des données »).

2. Il ne porte que sur les cohortes 2022-2024. Les cohortes 2019-2021, qui fournissent l'essentiel de l'information postérieure au traitement dans le modèle principal, ne sont pas testables faute d'antériorité suffisante : la fenêtre de données commence en 2017, ce qui ne laisse que deux années antérieures aux radars posés en 2019 et trois à ceux de 2020.

3. **L'écart placebo ne peut pas être retranché de l'estimation principale.** Les deux ne portent ni sur le même échantillon ni sur le même support temporel : la fenêtre placebo ne contient que la branche montante du profil pré-installation, quand le modèle principal contient la montée et le reflux. Toute arithmétique de débiaisage de ce type serait fautive, dans un sens comme dans l'autre.

Sur la mortalité, le placebo est sans portée : avec 174 tués dans les zones traitées de la fenêtre, son seuil de détection atteint 65 %. Il ne peut ni confirmer ni infirmer quoi que ce soit sur cet indicateur, et n'est mentionné ici que comme constat d'impuissance.

Périmètre du test. Le placebo a été conduit sur la fréquence des accidents et, à titre indicatif, sur les tués. Sur les blessés graves et les KSI, il ne figure pas dans les tableaux ci-dessus ; il a été appliqué dans le diagnostic de l'annexe 13.6, section E, à l'ensemble du groupe traité comme aux deux sous-groupes (spécification linéaire, g-2 et g-3, 500 m et 1 km) : sur l'ensemble, les douze configurations portant sur les tués, les blessés graves et les KSI s'écartent de zéro de 26,7 points en médiane, toutes dans le sens d'une baisse, quatre significativement, avec des seuils de détection propres de 28 % à 67 % à la baisse ; sur les sites nouveaux, l'écart est du même signe que sur la fréquence et plus marqué. Sur ces deux indicateurs, l'absence de biais de spécification n'est donc ni établie ni infirmée, et la réserve documentée ici s'y étend, même si la conclusion de non-détection qui les concerne ne dépend pas du test. Le placebo a aussi été appliqué aux deux autres estimateurs de l'étude, sur les mêmes cohortes et dans les quatre configurations principales (g-2 et g-3, 500 m et 1 km) : l'Empirical Bayes renvoie +12,5 %, +7,2 %, +12,8 % et +10,0 %, tous significatifs (écart médian 11,2 points), Callaway-Sant'Anna B +14,1 %, +4,3 %, +9,3 % et +5,5 % (un significatif, écart médian 7,4 points). Aucun des trois estimateurs ne restitue donc le zéro attendu, chacun avec un signe qui lui est propre (annexe 13.6, section E).

Diagnostic. Examiné à l'intérieur des cohortes, c'est-à-dire à composition constante sur les cinq années qui précèdent la pose, le profil d'accidentalité antérieur à la pose est monotone **croissant** et culmine l'année précédant l'installation (coefficients de -7 à -2 ans : -0,525 ; -0,399 ; -0,141 ; -0,122 ; -0,121 ; -0,027, référence à -1, estimés respectivement sur 111, 265, 414, 414, 414 et 414 zones : les points -7 et -6 ne reposent que sur les cohortes 2024, puis 2023 et 2024, la composition n'est constante qu'à partir de -5 ans, et la montée vaut 0,141 point de log sur ce seul segment). La trajectoire complète monte puis redescend : aucune tendance monotone ne peut l'épouser, et le signe du biais dépend de la branche qui domine la fenêtre d'estimation. Cela explique que le placebo, dont la fenêtre ne contient que la montée, produise un θ négatif, tandis que le modèle principal, qui contient les deux branches, produise un θ positif. Ce profil est celui des cohortes 2022-2024 (414 zones traitées sur 1 002). Le même profil estimé sur les cohortes 2019-2021 n'est pas concluant : ces cohortes ne disposent que de deux à quatre années antérieures, et l'année de référence de la cohorte 2021 est 2020, année du creux de circulation, ce qui rehausse mécaniquement les années plus anciennes ; les coefficients obtenus (+0,364 quatre ans avant la pose, $p = 0,031$, sur les 168 zones concernées ; -0,015 à trois ans ; +0,061 à deux ans) ne dessinent pas de profil monotone et ne peuvent pas être lus comme une trajectoire. Ils sont produits par l'étape 10 du programme de reproduction.