

Guide d'évaluation de la qualité des données d'un dispositif de surveillance épidémiologique en santé animale Version « bêta »

Mathilde PALUSSIÈRE
Anne BRONNER
Pascal HENDRIKX
Didier CALAVAS

Janvier 2014

L'analyse des données collectées dans le cadre d'un dispositif de surveillance en santé animale permet d'évaluer son fonctionnement ainsi que la situation épidémiologique du danger sanitaire surveillé. Toutefois, la qualité de cette analyse dépend fortement de la qualité des données recueillies ; l'évaluation de la qualité des données est donc un préalable indispensable à l'analyse et à l'interprétation des résultats de surveillance.

Ce guide propose une méthodologie d'évaluation de la qualité des données collectées dans le cadre des dispositifs de surveillance en santé animale, et gérées à l'aide du système d'information de la DGAL, SIGAL. Il comprend trois parties :

- Un glossaire et un recueil de définitions des termes spécifiques associés au fonctionnement des dispositifs de surveillance ainsi qu'à la qualité des données et à leur gestion dans SIGAL ;
- Une méthodologie d'évaluation illustrée à partir de l'évaluation de la qualité des données du dispositif de déclaration obligatoire des avortements chez les bovins en France ;
- Un formulaire d'évaluation à compléter, présenté en annexe (Annexe I).

La version du guide présentée ici a été élaborée au sein de l'Anses-Lyon, et a été utilisée pour évaluer le dispositif de déclaration des avortements chez les bovins. Toutefois, il s'agit d'une version « bêta » qui méritera très certainement d'être enrichie par les retours des différents évaluateurs qui seront amenés à utiliser ce guide.

Lorsqu'une partie du guide se rapporte à une section du formulaire à compléter, la section correspondant est précisée dans le titre de la partie. La mention « → **Compléter...** » indique le moment où la section correspondante du formulaire doit être renseignée.

Par ailleurs, chaque partie est illustrée par un exemple basé sur l'évaluation de la qualité des données du dispositif de déclaration obligatoire des avortements chez les bovins (DA). Les passages du guide qui correspondent à cet exemple sont identifiés comme suit :

Exemple d'évaluation : le dispositif de déclaration des avortements
<i>Contenu de l'exemple</i>

SOMMAIRE

GLOSSAIRE ET DEFINITIONS	5
INTRODUCTION	9
PARTIE 1 - ACTIONS PREALABLES A L'EVALUATION DE LA QUALITE DES DONNEES PROPREMENT DITE (SECTIONS 1 ET 2 DU FORMULAIRE)	11
1.1 CADRAGE DE L'EVALUATION (SOUS-SECTION 1-A DU FORMULAIRE)	11
1.2 DESCRIPTION DU DISPOSITIF (SOUS-SECTIONS 1-B A 1-D DU FORMULAIRE)	11
1.2.1 <i>Modalités de mise en œuvre du dispositif</i>	12
1.2.2 <i>Circulation des données</i>	16
1.2.3 <i>Modalités de gestion informatique des données</i>	19
1.3 SELECTION ET DESCRIPTION DES DONNEES (SECTION 2 DU FORMULAIRE)	22
1.3.1 <i>Sélection des données</i>	22
1.3.2 <i>Description et catégorisation des données</i>	24
PARTIE 2 - EVALUATION QUALITATIVE DE LA QUALITE DES DONNEES (SECTION 3 DU FORMULAIRE)	25
PARTIE 3 - EVALUATION QUANTITATIVE DE LA QUALITE DES DONNEES (SECTIONS 4 ET 5 DU FORMULAIRE)	28
3.1 EVALUATION DES SPECIFICATIONS GENERALES D'ENREGISTREMENT (SECTION 4 DU FORMULAIRE)	28
3.1.1 <i>Choix des critères d'évaluation</i>	28
3.1.2 <i>Construction et mesure des indicateurs</i>	29
3.2 EVALUATION QUANTITATIVE DE LA QUALITE DES DONNEES RESPECTANT LES SPECIFICATIONS GENERALES D'ENREGISTREMENT (SECTION 5 DU FORMULAIRE)	30
3.2.1 <i>A l'échelle des données (sous-section 5-A du formulaire)</i>	31
3.2.2 <i>A l'échelle de l'entité élémentaire de surveillance (sous-section 5-B du formulaire)</i>	34
3.2.3 <i>Etude de la cohérence (sous-section 5-C du formulaire)</i>	36
3.3 CONSTRUCTION ET MESURE D'INDICATEURS SYNTHETIQUES (SECTION 6 DU FORMULAIRE)	37
3.3.1 <i>Principe de construction des indicateurs</i>	38
3.3.2 <i>Fixation des valeurs cibles</i>	38
PARTIE 4 - PRESENTATION DES RESULTATS.....	40
4.1 RESULTATS DE L'EVALUATION QUALITATIVE	40
4.2 RESULTATS DE L'EVALUATION QUANTITATIVE.....	41
PARTIE 5 - RECHERCHE DE CORRELATIONS (SECTION 7 DU FORMULAIRE).....	50
PARTIE 6 - DISCUSSION ET INTERPRETATION DES RESULTATS	51
6.1 FACTEURS INFLUENÇANT LA QUALITE DES DONNEES	51
6.2 AXES D'AMELIORATION	58
PARTIE 7 - ACTIONS A METTRE EN PLACE APRES L'EVALUATION DE LA QUALITE DES DONNEES AU SENS STRICT	60
7.1 RETOUR D'INFORMATION	61
7.2 MISE EN ŒUVRE DES AXES D'AMELIORATION	61
7.3 SUIVI REGULIER DE LA QUALITE DES DONNEES	61
ANNEXE I – FORMULAIRE D'EVALUATION DE LA QUALITE DES DONNEES	64
ANNEXE II – EXEMPLE D'UNE FEUILLE DE COMMÉMORATIFS (HAUTE-SAVOIE, 74)	71
ANNEXE III – DESCRIPTION DES DONNEES A L'AIDE DE METADONNEES	72

GLOSSAIRE ET DEFINITIONS

Abréviations

BDNI	Base de données nationale de l'identification du ministère de l'agriculture, de l'agroalimentaire et de la forêt
DA	Déclaration d'avortement
EAT	Epreuve à l'antigène tamponné
EBRUBVA	Plan d'analyse Elevage - Brucellose bovine – Avortement
ELISA	Enzyme-linked immunosorbent assay
FC	Fixation du complément
iELISA	ELISA indirecte
IS	Indicateur synthétique
MAM	Triplet matrice – analyte – méthode
NS	Note de service
EES	Entité élémentaire de surveillance

Acteurs

ADILVA	Association française des directeurs et cadres de laboratoires vétérinaires publics d'analyses
Anses	Agence nationale de sécurité sanitaire – Alimentation, santé, travail
CoSiR	Coordinateur SIGAL régional
CSD-ESA	Centre de service des données pour l'épidémiologie en santé animale
DGAL	Direction générale de l'alimentation, du Ministère de l'agriculture, de l'agroalimentaire et de la forêt
DDecPP	Direction départementale en charge de la protection des populations
EDE	Etablissement départemental de l'élevage
GDS France	Fédération nationale des groupements de défense sanitaire
GDS	Groupement de défense sanitaire
GTV	Groupement technique vétérinaire
LNR	Laboratoire national de référence
LVD	Laboratoire vétérinaire départemental
SNGTV	Syndicat national des groupements techniques vétérinaires
VS	Vétérinaire sanitaire
Plateforme ESA	Plateforme d'épidémiologie en santé animale

Dispositifs de surveillance

Surveillance active

Elle repose sur la recherche de données par des actions programmées à l'avance. Elle peut être exhaustive, sur échantillon représentatif ou sur échantillon orienté (Hoinville et al., in press).

Surveillance évènementielle

Elle repose sur la production spontanée de données de surveillance. Cela peut être la déclaration spontanée de cas ou suspicions de cas de la maladie surveillée ou l'utilisation de données déjà existantes (Hoinville et al., in press).

Qualité des données

Donnée Elle est définie comme un fait, une valeur affectée à une variable (les notions de donnée et de variable seront ici indifféremment associées à l'appellation « donnée »).

Donnée de qualité élevée

C'est une donnée qui *est convenable pour l'usage que l'on souhaite en faire* (« *fitness for use* »). Ainsi, cette notion est étroitement dépendante des utilisateurs de la donnée et des personnes conduisant l'évaluation de la qualité des données.

Concept de qualité des données

La qualité d'une donnée est un concept multidimensionnel. Certaines dimensions se rapportent aux objectifs globaux de qualité des données habituellement rencontrés. D'autres se rapportent à des facteurs qui dépendent strictement de la perception qu'ont les utilisateurs des données : elles peuvent influencer les dimensions citées précédemment.

Complétude Absence de valeur nulle.

Format Absence d'erreur de syntaxe (orthographe ou format).

Validité Concordance entre la donnée collectée et la réalité ou une liste de valeurs possibles.

Vitesse de circulation des données

Délai entre la date de l'évènement ayant généré la donnée et la date à partir de laquelle la donnée est disponible.

Cohérence Cohérence de plusieurs données entre elles.

Compréhension

Facilité de compréhension de la donnée sur sa signification et sur ce qui est attendu en termes de format.

Pertinence Utilité de la donnée, qui concorde avec les besoins des utilisateurs.

Crédibilité Perception de l'exactitude de la donnée par l'utilisateur.

Accès / utilisation

Facilité d'accès et d'utilisation, à un niveau pratique (rapide, sans intervention de personnes extérieures).

Systemes d'information de la DGAL et des laboratoires

SIGAL	Systeme d'information de la DGAL.
Lims	Laboratory information management system : systeme d'information d'un laboratoire. Il est utilise par les differents laboratoires en France pour la saisie et l'envoi des donnees dans SIGAL.
EDI	Echange de donnees informatise. Il permet a deux partenaires de communiquer quels que soient leurs systemes informatiques respectifs. L'EDI Sacha est l'EDI des laboratoires, qui leur permet de communiquer avec leurs differents partenaires, y compris l'Etat.
Etablissement	Il est caracterise par plusieurs types d'identifiant dont au moins un numero SIRET et un numero SIREN. Il peut detenir un ou plusieurs ateliers selon le type de production de l'elevage (ex : atelier de bovins allaitants). Lorsque l'etablissement correspond a une exploitation bovine, il comporte au moins un type d'identifiant supplementaire : un numero EDE.
Atelier	Il caracterise le type de production de l'elevage. Un atelier bovin peut etre laitier ou allaitant. Deux ateliers d'un meme etablissement peuvent avoir des numeros d'EDE differents lorsqu'ils sont localises sur des communes differentes.
Acte de reference	Plusieurs actes de reference sont defines dans SIGAL. Ils representent toutes les actions qui peuvent etre geres dans SIGAL. Par exemple : une declaration d'avortement.
Intervention	Elle reunit un « quoi », l'acte de reference, un « qui », l'acteur qui a realise l'acte, un « ou », l'etablissement ou a eu lieu l'acte, et un « quand », date a laquelle a eu lieu l'acte. Une intervention peut etre programmee ou non programmee. Normalement, pour chaque acte realise une intervention est creee.
INP	C'est une intervention non programmee. Elle est creee dans SIGAL une fois que l'acte a ete realise, l'acte en question ne pouvant pas etre prevu.
IP	C'est une intervention programmee. Elle est creee dans SIGAL avant que l'acte ne soit realise, avec un « quoi », un « ou », une date previsionnelle (au plus tot et/ou de preference et/ou au plus tard) et eventuellement un « qui ». Sont prevus en plus dans cette IP les analyses demandees et les animaux concernes par ces analyses.
DAI	C'est une demande d'analyse informatisee. Elle n'existe que dans le cadre des IP, et elle est editee a partir de cette IP lors de l'impression d'un DAP. Elle est alors envoyee au laboratoire auquel les analyses sont demandees. Elle contient des commemoatifs (le « quoi », le « ou » et eventuellement le « qui ») ainsi que les analyses demandees, en premiere et en deuxieme intention.
DAP	Document d'accompagnement des prelevements. Il est edite et imprime a partir d'une IP, puis transmis au detenteur des animaux concernes ou au maitre d'oeuvre (personne qui doit realiser l'acte, par exemple le VS). Il reprend des commemoatifs (le « quoi », le « ou » et eventuellement le « qui »), les analyses demandees, les prelevements a realiser et le laboratoire ou ils doivent etre envoyes. Il doit etre

complété lors de la réalisation des prélèvements demandés. Il est ensuite envoyé au laboratoire avec les prélèvements.

- RAI** Résultat d'analyses informatisé. Ce fichier informatique contient les données relatives aux analyses réalisées par le laboratoire et au moins une partie des commémoratifs. Il est créé par le laboratoire et exporté dans SIGAL.
- MCD** Modèle conceptuel des données. Il permet de représenter de façon structurée et formalisée des données ainsi que les types de relations existants entre ces données.
- SRD** Schéma relationnel des données. C'est un modèle conceptuel des données adapté aux bases de données relationnelles.

Divers

- EES** Entité élémentaire de surveillance. Elle est caractérisée par :
- la date de la réalisation de l'acte ;
 - le lieu de réalisation de l'acte ;
 - pour la surveillance événementielle le ou les animaux faisant l'objet de la suspicion clinique, et pour la surveillance active celui ou ceux faisant l'objet des prélèvements.
- Exemples d'EES pour trois dispositifs de surveillance :
- dispositif de DA : une déclaration avortement, c'est-à-dire le triplet bovin / date de visite / numéro EDE ;
 - salmonelles aviaires : le triplet identifiant du bâtiment / date du prélèvement / identifiant de l'atelier ;
 - surveillance active de la brucellose : le triplet date de visite pour tout ou partie de la surveillance active / numéro EDE / identifiants des animaux prélevés.

INTRODUCTION

Validation de la démarche

La mise en place d'une démarche d'évaluation et plus globalement d'amélioration de la qualité des données est une décision importante, à prendre en accord avec tous les acteurs concernés, que ce soit en matière d'animation, de collecte, de saisie, de traitement et d'analyse des données. En effet, la réussite d'un tel projet dépend directement de l'implication des acteurs dans les différentes étapes.

En particulier, les avantages et les difficultés à mener une telle démarche doivent être discutés :

- Améliorer la qualité des données permet d'assurer la fiabilité de l'analyse des résultats de surveillance ainsi que la fiabilité des indicateurs de fonctionnement (German et al., 2001) ;
- A l'inverse, cette démarche implique un investissement en temps, immédiat et limité (dans le cadre de l'évaluation) mais également à plus long terme. En effet, une évaluation de la qualité des données ne doit pas être une action isolée : le plan d'action pour l'amélioration, le suivi, la vérification de la mise en œuvre et les adaptations mises en place doivent former un tout.

Constitution de l'équipe d'évaluation

Une équipe responsable de l'évaluation de la qualité des données doit être constituée. Elle comprend naturellement l'animateur du dispositif étudié. La ou les personnes (nombre limité, idéalement deux personnes et maximum trois ou quatre personnes) (Hendrikx et al., 2011) devant mener concrètement l'évaluation de la qualité des données doivent avoir des bases dans les domaines suivants :

- Fonctionnement général des dispositifs de surveillance et de SIGAL,
- Acteurs impliqués dans les dispositifs de surveillance,
- Analyse et traitement des données (à un niveau uniquement descriptif).

L'évaluation peut être interne ou externe au dispositif, mais dans les deux cas l'encadrement de l'évaluation doit être clairement défini. Dans un souci d'objectivité et d'impartialité, il est intéressant que la ou les personnes menant l'évaluation soient externes au dispositif, c'est-à-dire qu'elles n'occupent pas de fonction dans le pilotage, l'animation ou la coordination. Afin d'éviter tout conflit d'intérêt, elles ne doivent pas être salariées du dispositif lui-même ou de l'un de ses partenaires.

L'équipe d'évaluation sera en charge de la conduite de l'évaluation de la qualité des données et de la communication qui devra y être associée, en préalable et suite à l'évaluation. Cette équipe doit être clairement identifiée mais également clairement identifiable par les différents acteurs impliqués directement ou indirectement dans le dispositif.

D'une part, il faut communiquer autour du principe de la démarche, sur ce qui va être fait : des mesures de la qualité, des entretiens avec les acteurs, éventuellement un planning prévisionnel, etc. D'autre part, il est important de communiquer avec les personnes ne se trouvant pas directement dans l'évaluation sur l'intérêt d'avoir des données de bonne qualité : fiabilité de l'analyse des résultats de surveillance, fiabilité des indicateurs de fonctionnement (German et al., 2001).

La même équipe ou une autre sera, quant à elle, en charge du suivi de la qualité des données après la première évaluation. Cette équipe est à désigner au même moment que l'équipe en charge de l'évaluation, c'est-à-dire dès le commencement. Assigner le suivi dans le temps à une personne ou un groupe de personnes permet de limiter les risques d'abandon de la « démarche d'évaluation de la qualité des données », de clarifier le message pour les parties prenantes et d'asseoir l'importance de la démarche dans le temps.

Matériel nécessaire à la conduite de l'évaluation

La conduite de la méthodologie proposée dans ce guide nécessite de recueillir préalablement les différents textes réglementaires et non réglementaires relatifs au dispositif de surveillance étudié. Par ailleurs, il faut également obtenir une extraction des données enregistrées dans SIGAL dans le cadre de la mise en œuvre du dispositif étudié.

Principe de la méthodologie proposée dans ce guide

La méthodologie proposée ici peut être divisée en quatre séquences :

- une étape préalable à la mesure de la qualité, comprenant notamment une description du dispositif et de son fonctionnement (partie A.) ;
- la mesure de la qualité (parties B et C) ;
- l'exploitation des résultats de la mesure de la qualité : elle permet d'identifier des facteurs influençant la qualité des données et de faire ressortir des axes d'amélioration (parties D, E et F) ;
- une étape postérieure à l'évaluation de la qualité, comprenant notamment le suivi (partie G).

La mesure de la qualité des données consiste globalement à calculer des indicateurs relatifs à différentes dimensions de la qualité (cf. glossaire). Elle est conduite aux niveaux national et départemental, incluant tous les départements où la mise en œuvre du dispositif de surveillance a été effective durant la période étudiée. Elle porte sur les données à visée de surveillance (les données de gestion ne sont pas étudiées).

L'évaluation est qualitative et quantitative.

La partie qualitative est menée à l'aide d'entretiens auprès de différents acteurs dans plusieurs départements. Elle apporte une aide à la compréhension et à l'interprétation des résultats de l'évaluation quantitative.

Le volet quantitatif est développé en trois étapes. La première étape est réalisée sur tous les enregistrements disponibles pour la période d'étude. Elle correspond à l'évaluation du respect de spécifications générales d'enregistrement, s'appuyant sur la définition de critères permettant d'élaborer des indicateurs. La deuxième étape se focalise sur la qualité des données associées aux enregistrements respectant ces spécifications générales. Des indicateurs sont construits en rapport avec les dimensions de la qualité sélectionnées. Cette étape est réalisée à l'échelle des données puis à l'échelle des entités élémentaires de surveillance (EES). La dernière étape correspond à la construction d'indicateurs synthétiques permettant d'avoir une vision globale de la qualité des données. Ce volet a pour but de disposer d'indicateurs permettant d'identifier les données pour lesquelles la qualité est insuffisante (données qui ne reflètent pas la réalité ou qui ne peuvent pas être traitées) et les facteurs explicatifs.

La méthodologie d'évaluation présente en détail les étapes de l'évaluation, et propose des indicateurs dont l'étude apparaît nécessaire dans tous les cas. D'autres indicateurs peuvent être évalués en fonction du choix de l'évaluateur, de l'intérêt d'étudier ces indicateurs au vu des objectifs de l'étude et des moyens pratiques disponibles (notamment matériels, humains et financiers). Pour permettre une interprétation et une analyse des résultats réaliste et fiable, le nombre d'indicateurs utilisés pour mesurer la qualité des données ne doit pas être trop important. Il est nécessaire de concilier deux points potentiellement opposés : la finesse de l'évaluation (beaucoup d'indicateurs, avec chaque indicateur dépendant d'un unique facteur influençant la qualité des données) et le réalisme (moyens pratiques disponibles, possibilité de sélectionner de tels indicateurs, niveau de détail de l'évaluation réellement nécessaire).

ACTIONS PREALABLES A L'EVALUATION DE LA QUALITE DES DONNEES PROPREMENT DITE (SECTIONS 1 ET 2 DU FORMULAIRE)

1.1 Cadrage de l'évaluation (sous-section 1-A du formulaire)

Le cadrage de l'évaluation doit permettre de préciser :

- l'objet de l'évaluation, c'est-à-dire le dispositif concerné, et s'il est concerné dans son intégralité ou uniquement en partie ;
- l'objectif principal de l'évaluation de la qualité des données : cela peut être la réalisation d'un état des lieux, la comparaison avec un travail antérieur ou l'identification d'éléments d'aide à la prise de décisions pour une évolution du dispositif ;
- les différents types d'acteurs concernés par l'évaluation. Cette liste d'acteurs pourra éventuellement être complétée, en particulier suite à la description de la circulation des données.

→ Compléter la sous-section 1-A – Cadrage de l'évaluation

Exemple d'évaluation : le dispositif de déclaration des avortements
<p><u>Objectifs et contexte de l'évaluation</u></p> <p>Les objectifs de l'étude sont d'évaluer la qualité des données issues des déclarations d'avortements chez les bovins (DA), d'identifier les facteurs influençant cette qualité et enfin de proposer des pistes d'amélioration.</p> <p>En effet, le dispositif de DA n'a pas évolué depuis sa mise en place en 1965, et il présente une faible sensibilité (parmi les éleveurs qui détectent un avortement, seul un sur quatre le déclarerait) (Bronner et al., 2013). Une évolution de ce dispositif est donc nécessaire. Ce besoin a été relayé par les différents acteurs concernés par ce dispositif, réunis depuis 2011 au sein du groupe de suivi « Surveillance des maladies abortives d'intérêt pour l'Etat en élevage de ruminants » de la Plateforme nationale de surveillance épidémiologique en santé animale (Plateforme ESA). Pour mettre en place cette évolution, une évaluation approfondie de son fonctionnement et de la qualité des données collectées est requise.</p>

1.2 Description du dispositif (sous-sections 1-B à 1-D du formulaire)

1.2.1 Modalités de mise en œuvre du dispositif

La description du dispositif étudié doit comprendre des informations sur :

- l'objet de la surveillance, c'est-à-dire la maladie surveillée ;
- les objectifs du dispositif ;
- la maîtrise d'ouvrage et l'animation du dispositif ;
- les modalités de surveillance (active ou événementielle, population surveillée) ;
- les modalités d'échantillonnage (si nécessaire) ;
- les modalités de réalisation de l'acte (réalisation de la visite, prélèvements réalisés) ;
- les modalités d'analyse ;
- les modalités de gestion des données ;
- les modalités de formation des acteurs et de communication autour du dispositif ;
- les modalités d'évaluation du dispositif ;
- l'EES considérée pour l'étude, et si l'étude se fait par animal ou par groupes d'animaux¹ ;
- les documents disponibles encadrant la surveillance et la collecte des données, qu'ils soient réglementaires (arrêtés, notes de service, ordres de méthode...) ou non (articles, fiches de plan).

→ Compléter la sous-section 1-B – Description de la mise en œuvre du dispositif

Exemple d'évaluation : le dispositif de déclaration des avortements

Objet de la surveillance

La brucellose bovine est une maladie infectieuse, contagieuse et zoonotique dont la principale manifestation clinique chez les bovins est l'avortement, particulièrement à partir de cinq mois de gestation. Elle a été progressivement éradiquée en France par des campagnes de prophylaxie généralisée mises en place dans les années 1960. Le statut officiellement indemne en brucellose bovine a été accordé à la France en 2005 par la Commission européenne (Anonyme, 2005a), après une absence de foyer autochtone depuis 2003. La découverte d'un foyer autochtone en 2012² n'a pas entraîné de retrait de ce statut compte tenu de sa gestion satisfaisante.

Objectifs du dispositif de DA

La surveillance de la brucellose bovine a pour objectifs de garantir le maintien du statut sanitaire favorable du pays et d'assurer une détection précoce de toute réapparition d'un foyer sur le territoire. Elle repose sur une surveillance active (dépistage annuel des élevages de bovins), ainsi que sur une surveillance événementielle, la DA.

Par ailleurs, depuis 2012, le dispositif de DA est aussi le support d'une surveillance événementielle de la Fièvre Q³ qui a été mise en place dans dix départements pilotes (2012a).

¹ Par exemple, on parle généralement d'animaux pour les bovins et de groupes d'animaux pour les porcins ou les volailles.

² Mise en évidence d'un foyer de brucellose en avril 2012 en Haute-Savoie.

³ La Fièvre Q est une maladie infectieuse abortive due à *Coxiella burnetii* à laquelle les ruminants sont sensibles. C'est également une zoonose.

Maîtrise d'ouvrage et animation du dispositif

L'animation et la maîtrise d'ouvrage sont sous la responsabilité de l'Etat. La DGAL et à l'échelle locale les DDecPP sont donc responsables de la mise en œuvre et du suivi du dispositif de DA, définis par l'arrêté ministériel du 22 avril 2008 fixant les mesures techniques et administratives relatives à la prophylaxie collective et à la police sanitaire de la brucellose des bovinés, ainsi que par plusieurs notes de service (2010b ; 2010a). Par ailleurs, les GDS peuvent se voir confier une partie de la gestion des DA, par délégation de service public (Anonyme, 2011b). Lors de la campagne 2011-2012 (du 31 août au 1^{er} juillet), cette délégation a été effective dans huit départements. Dans deux autres départements, la gestion était partagée entre la DDecPP et le GDS⁴, ces départements ayant probablement débuté une délégation au cours de cette campagne.

La prise en charge financière de la visite du vétérinaire et des analyses réalisées dans le cadre de ce dispositif revient à l'Etat.

Le dispositif de surveillance est animé au niveau national par la DGAL (Bureau de la santé animale), avec l'appui du groupe de suivi « Surveillance des maladies abortives d'intérêt pour l'Etat en élevage de ruminants » de la Plateforme ESA. Le laboratoire national de référence de la brucellose (LNR, situé à l'Anses – Maisons Alfort) a la charge de l'animation du réseau des laboratoires agréés⁵.

Modalités de surveillance

La DA est un dispositif de surveillance événementielle visant à surveiller les avortements survenant au sein de la population des femelles gestantes sur l'ensemble du territoire français. Ainsi, tout éleveur observant un avortement ou ses signes cliniques dans son élevage est tenu d'en informer immédiatement son VS.

Modalités d'échantillonnage

Le dispositif de DA est un dispositif de surveillance qui se veut exhaustif : tout avortement doit être déclaré.

Modalités de réalisation de l'acte

Un avortement est défini comme suit : « Est considéré comme avortement l'expulsion d'un fœtus ou du veau, soit né mort, soit succombant dans les 48 heures après la naissance » (en vigueur du 7 août 2003 au 2 juillet 2012). Le VS, mandaté par l'Etat pour réaliser des opérations sanitaires en élevage pour le compte de ce dernier, réalise une visite dans l'exploitation où est détenue la femelle ayant avorté. Au cours de cette visite, il recueille des commémoratifs et réalise des prélèvements : une prise de sang systématique, ainsi qu'un écouvillon endocervical (au niveau de col de l'utérus) dans certains départements. La feuille de commémoratifs et les prélèvements sont envoyés à un laboratoire agréé pour la sérologie et éventuellement la bactériologie vis-à-vis de la brucellose : cet envoi fait également office de DA. Il est conseillé d'envoyer l'écouvillon le plus rapidement

⁴ Dans les départements de l'Eure (27) et de la Manche (50), sur la campagne étudiée, un peu plus de 40% des DA ont été gérées par la DDecPP et 60% par le GDS.

⁵ La liste des laboratoires agréés pour les différentes méthodes de diagnostic de la brucellose est disponible sur le site internet http://agriculture.gouv.fr/IMG/pdf/Brucellos-Liste_des_laboratoires_agrees_pour_les_différentes_méthodes_de_diagnostic_V17_cle06c691.pdf

possible (moins de 24h si possible selon la NS 2008-8322 relative à la surveillance de la brucellose (2008b)).

Le dispositif de DA a entraîné en 2011 la réalisation de 61 631 interventions pour 61 707 avortements, ce qui représente une baisse de 1,4% par rapport à l'année précédente. Ces interventions ont été réalisées dans 34 707 exploitations différentes (baisse de 2%), ce qui représente à peine plus de 16% des élevages soumis à la prophylaxie (Anses, 2013).

Modalités d'analyse

Le laboratoire agréé réalise une analyse sérologique sur la prise de sang, dans un premier temps en effectuant soit une iELISA (indirect Enzyme-linked immunosorbent assay), soit une EAT (épreuve à l'antigène tamponné). Si l'analyse réalisée est positive, le laboratoire réalise alors une FC (fixation du complément), méthode plus spécifique que les précédentes.

Lorsque l'analyse sérologique est confirmée positive en FC et que l'écouvillon est disponible, une culture et une identification bactérienne sont conduites sur cet écouvillon par un laboratoire agréé pour la bactériologie (2010b ; 2010a), qui peut être le laboratoire ayant réalisé la sérologie ou un autre laboratoire (qui a reçu l'écouvillon transmis par le premier laboratoire). Si l'écouvillon n'est pas disponible, le vétérinaire refait une visite pour effectuer ce prélèvement, à condition que ce soit dans les quinze jours suivant l'avortement et qu'aucun traitement antibiotique n'ait été fait (2010b ; 2010a). Sinon, un abattage diagnostique doit être réalisé à des fins de cultures bactériennes, réalisées sur au moins trois paires de nœuds lymphatiques (les nœuds lymphatiques rétro-pharyngiens, rétro-mammaires, utérins et lombo-aortiques) (2010b ; 2010a).

Parmi la totalité des avortements déclarés en 2011 (61 707), seuls onze ont donné un résultat sérologique confirmé positif (EAT ou iELISA positive et FC positive), soit 0,018 %. Les recherches bactériologiques réalisées consécutivement par les laboratoires agréés pour le diagnostic direct n'ont conduit à aucun isolement de *Brucella* (Anses, 2013).

Modalités de gestion des données

La saisie et la centralisation des données dans SIGAL ont débuté en 2004. La saisie peut être réalisée par des agents de la DDecPP ou du GDS (par délégation de service public), et/ou par le laboratoire d'analyses vétérinaires, ce qui est le plus fréquent.

Les données saisies sont extraites régulièrement au niveau départemental par les DDecPP afin d'assurer le paiement des VS ayant réalisé les visites, réalisé par l'intermédiaire de Chorus⁶ depuis 2012.

Des extractions de ces données au niveau national sont actuellement possibles sur demande spécifique au Ministère chargé de l'agriculture. Les données seront prochainement disponibles sur simple requête via le Centre de service des données d'épidémiosurveillance en santé animale (CSD-ESA), en cours de mise en place par le Ministère dans le cadre de la Plateforme ESA. Les données sont analysées une fois par an par le Bureau de la santé animale de la DGAL (Anses).

⁶ Chorus est un logiciel qui permet de gérer les dépenses, les recettes non fiscales et la comptabilité de l'État en France dans le cadre de la loi organique relative aux lois de finances.

Les droits d'écriture et de lecture conditionnent fortement la gestion des données. Les droits d'écriture sont partagés entre les DDecPP et les GDS pour les données concernant les exploitations de leur département. Par ailleurs, il leur est possible de créer une intervention pour un autre département. En ce qui concerne les droits de modification / suppression, telle que par exemple la suppression d'une DA enregistrée dans SIGAL, ils sont définis localement par l'administrateur SIGAL départemental. Enfin, pour les droits de lecture, les DDecPP et les GDS ont accès aux données de leur département, la DGAL et GDS France ont accès aux données nationales et les VS ont accès aux données des éleveurs pour lesquels ils sont VS.

Modalités de formation et de communication

En 2011, dans le cadre de la formation continue au mandat sanitaire, les VS des dix départements pilotes pour la mise en place d'une surveillance de la fièvre Q ont suivi une session de formation sur le thème « Surveillance des avortements des ruminants : au-delà de la brucellose ». En 2013, une formation sur la DA a été proposée à tous les départements, et une quarantaine de sessions de formation sont prévues (Anonyme, 2011a).

Les résultats annuels de la surveillance de la brucellose font l'objet d'une information à l'attention des acteurs de terrain dans un numéro spécial « Maladies réglementées et émergentes » du Bulletin épidémiologique santé animale – alimentation Anses / DGAL (Anses). Localement, les DDecPP, GDS et GTV peuvent mener des actions d'information et de sensibilisation à l'attention des éleveurs et/ou des VS, que ce soit lors de réunions (réunions annuelles consacrées à la prophylaxie avec les vétérinaires, assemblées générales), par voie informatique (mail, site internet) ou postale (brochures).

Modalités d'évaluation du dispositif

Le dispositif de DA ne dispose actuellement pas d'indicateurs de fonctionnement qui pourraient permettre de le suivre régulièrement. Cependant, dans le cadre du groupe de travail « Surveillance des maladies abortives d'intérêt pour l'Etat en élevage de ruminants » de la Plateforme ESA, de tels indicateurs sont prévus.

L'évaluation du dispositif a été initiée en 2012, en se focalisant dans un premier temps sur l'estimation de la sensibilité et de la représentativité du dispositif actuel et l'étude du processus de décision conduisant les éleveurs et les VS à participer, ou non, au dispositif. Cette évaluation a montré que les avortements sont largement sous déclarés : environ trois quarts des éleveurs qui détecteraient un avortement ne le signaleraient pas, cette proportion étant plus importante pour les éleveurs allaitants que laitiers (Bronner et al., 2013). L'analyse de la qualité des données collectées s'inscrit à la suite de ces travaux.

EES

L'entité élémentaire de surveillance considérée est la déclaration d'avortement, c'est-à-dire le triplet identifiant EDE / date de la visite / numéro d'identification du bovin. L'étude se fait par animal, c'est-à-dire par bovin ayant avorté.

Documents disponibles

Il n'existe actuellement pas d'ordre de méthode national (c'est-à-dire de consignes officielles) présentant les modalités de collecte de saisie ou de centralisation dans SIGAL des données collectées dans le cadre du dispositif de DA chez les bovins. Réglementairement, ces modalités reposent uniquement sur des textes généralistes (Anonyme, 2007) ou sur des documents très spécifiques (comme la fiche de plan d'analyse). D'autres textes réglementaires précisent quelques points très précis (2009b ; 2010b ; 2010a ; 2012a). Par ailleurs, il existe des accords locaux entre les DDecPP et les laboratoires ainsi qu'entre les DDecPP et les GDS dans les départements où une délégation est mise en place, reposant sur une convention entre les différentes parties. Ces accords sont encadrés par la loi, mais sont adaptables et adaptés au sein de chaque département (Anonyme, 2011b).

1.2.2 Circulation des données

La description de la circulation des données permet d'avoir une meilleure connaissance de la vie des données, et d'une partie au moins des facteurs pouvant influencer leur qualité. Dans le cas des IP comme des INP, un schéma général de circulation des données peut être proposé (fig A, fig B).

Dans le cas des IP (fig A), une intervention est créée dans SIGAL par la DDecPP ou le GDS. A partir de cette intervention, une DAI est envoyée au laboratoire et un DAP est édité puis envoyé au détenteur des animaux ou à l'acteur qui réalise les prélèvements. Ce DAP est complété pendant la visite, transmis au laboratoire puis éventuellement à la DDecPP ou au GDS. Une fois les analyses réalisées, les données sont enregistrées par le laboratoire dans son Lims, puis exportées dans SIGAL via un RAI. Il est alors rattaché automatiquement à l'IP initialement créée.

Les différents acteurs ont accès aux données renseignées dans les interventions.

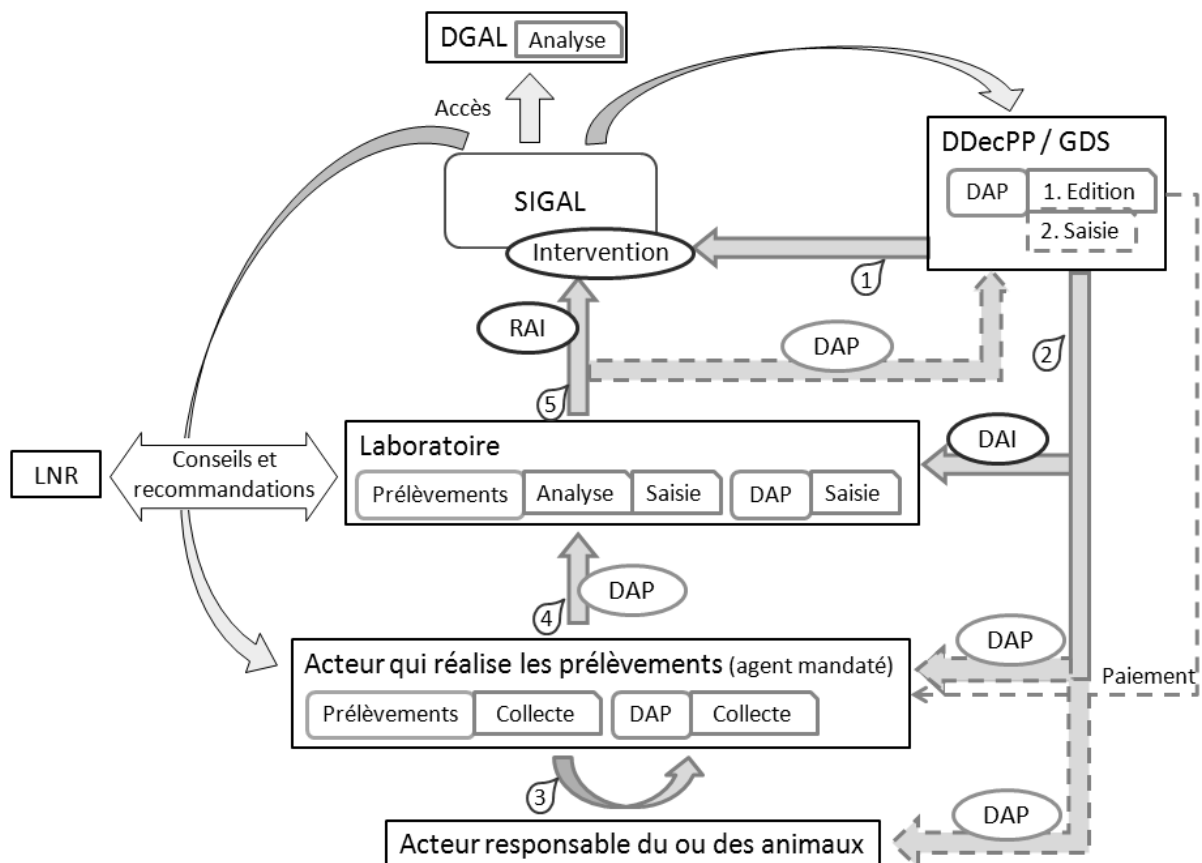


Figure A : Schéma de circulation des données pour les interventions programmées

Dans le cas des INP (fig B), la déclaration de l'évènement déclenche la réalisation d'une visite lors de laquelle les prélèvements sont réalisés et des données sont collectées sur une fiche de commémoratifs. Les prélèvements et la fiche de commémoratifs sont transmis au laboratoire. Une fois les analyses réalisées, les données sont enregistrées par le laboratoire dans son Lims, puis exportées vers une interface SIGAL – EDI Sacha via un RAI. Le RAI est alors utilisé par la DDecPP ou le GDS pour créer une intervention (qui est donc une INP) dans SIGAL.

Les différents acteurs ont accès aux données renseignées dans les interventions.

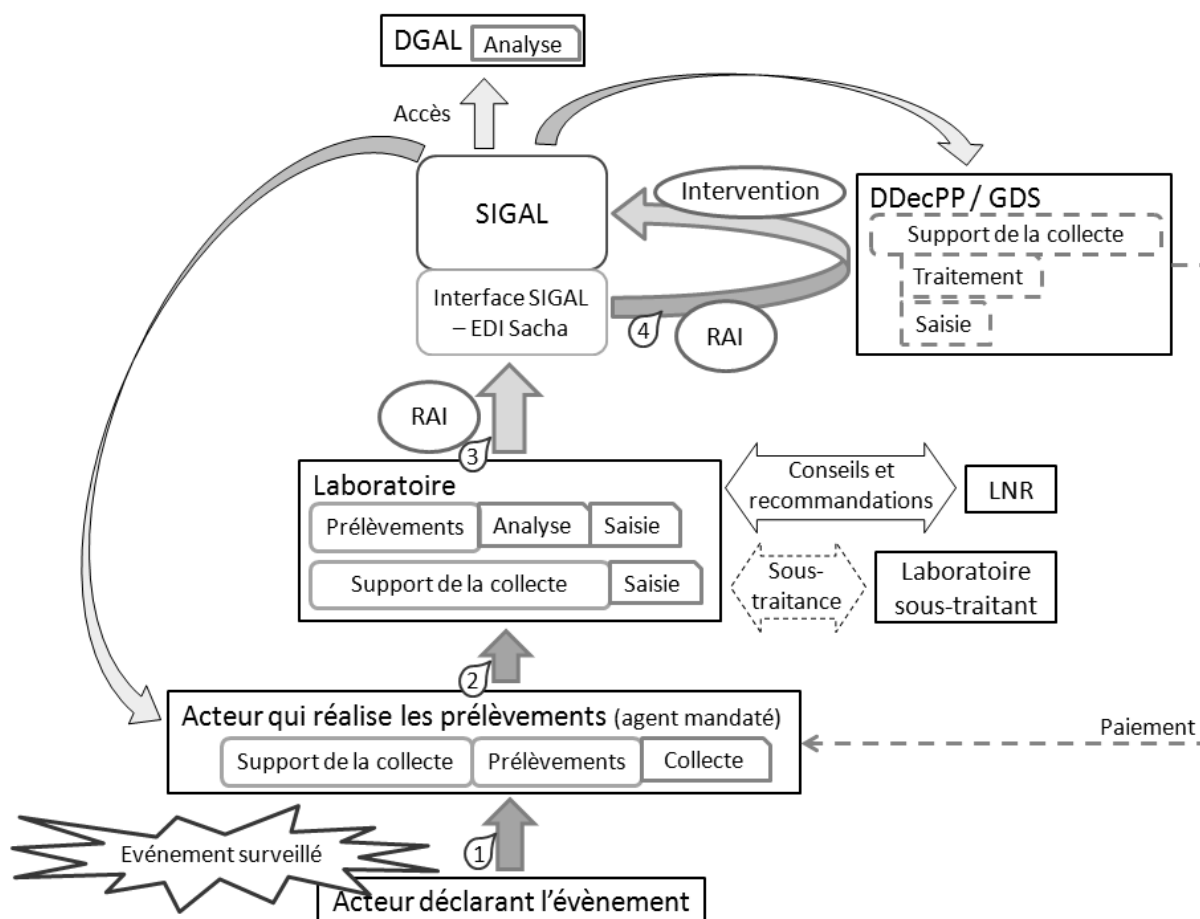


Figure B : Schéma de circulation des données pour les interventions non programmées

Certaines informations doivent être précisées sur ces schémas en fonction du dispositif étudié, notamment :

Pour les IP :

- la personne responsable des animaux faisant l'objet de l'intervention ;
- la personne qui réalise les prélèvements, en précisant si les données récupérées dans SIGAL sont utilisées par la DDecPP ou le GDS pour sa rémunération ;
- le LNR (ou équivalent) ;
- le rôle de la DDecPP et / ou du GDS dans la saisie des données dans SIGAL ;
- le destinataire du DAP, depuis son édition par la DDecPP ou le GDS jusqu'à la personne qui le conserve.

Pour les INP :

- l'évènement surveillé ;
- la personne déclarant l'évènement ;
- la personne qui réalise les prélèvements, en précisant si les données récupérées dans SIGAL sont utilisées par la DDecPP ou le GDS pour sa rémunération ;
- le support de la collecte des données (fichier informatique, document papier ou autre), et l'existence éventuelle d'une transmission de ce document par le laboratoire à la DDecPP ou au GDS ;
- le rôle de la DDecPP ou du GDS dans la saisie des données dans SIGAL ;
- le LNR (ou équivalent) ;

- la possibilité d'une sous-traitance des analyses à un autre laboratoire que le destinataire des prélèvements.

→ Compléter la sous-section 1-C – Description du flux des données

Exemple d'évaluation : le dispositif de déclaration des avortements
<p>Pour le dispositif de DA :</p> <ul style="list-style-type: none"> - <u>l'évènement surveillé</u> : tout avortement chez un bovin. Un avortement est défini comme suit : « <i>Est considéré comme avortement l'expulsion d'un fœtus ou du veau, soit né mort, soit succombant dans les 48 heures après la naissance</i> » ; - <u>la personne déclarant l'évènement</u> : la personne responsable de la femelle ayant avorté (généralement l'éleveur) ; - <u>la personne qui réalise les prélèvements</u> : le VS de l'éleveur. Il est intégralement rémunéré par l'Etat ; - <u>le support de la collecte des données</u> : la feuille de commémoratif (papier). Chaque département a sa propre feuille de commémoratifs (un exemple est présenté annexe II). Selon les départements, elle peut ou non être transmise à la DDecPP et/ou au GDS ; - <u>le rôle de la DDecPP ou du GDS dans la saisie des données</u> : selon les départements, la DDecPP et/ou le GDS peuvent saisir des données dans SIGAL ; - <u>le LNR</u> : Anses – Maison-Alfort ; - <u>la possibilité d'une sous-traitance des analyses</u> : la sérologie et la bactériologie peuvent être sous-traitées si le laboratoire qui reçoit les prélèvements n'est pas agréé pour l'une ou l'autre de ces analyses.

1.2.3 Modalités de gestion informatique des données

Le schéma relationnel des données (SRD) type permet de préciser les relations contraintes informatiquement dans SIGAL, ainsi que les attentes en matière d'enregistrement des données. Le SRD proposé ici (fig C) est applicable à tous les dispositifs dont les données sont gérées dans SIGAL.

Le SRD présente les relations et le type de relation entre entités : ce sont les chiffres reliant plusieurs rectangles. Par exemple, une occurrence de l'entité RAI peut être reliée à une ou à plusieurs occurrences de l'entité Analyse (1,n) alors qu'une occurrence de l'entité Analyse n'est reliée qu'à une et une seule occurrence de l'entité RAI (1,1).

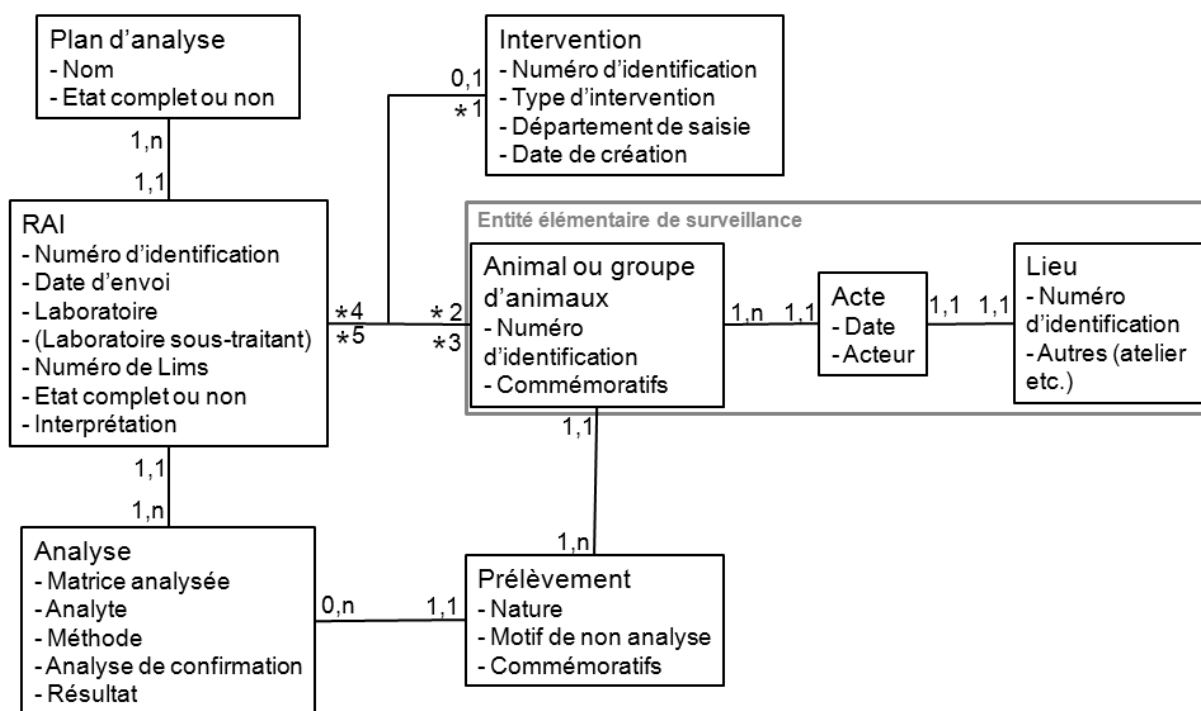


Figure C : Schéma relationnel des données

Ce schéma présente deux types de relations :

- des relations pour lesquelles la contrainte informatique imposée par SIGAL semble suffisante : il n'y a aucun intérêt à imposer des contraintes supplémentaires, voire ces contraintes supplémentaires seraient problématiques (par exemple, il est nécessaire de pouvoir renseigner plusieurs analyses dans un même RAI) ;
- des relations pour lesquelles la contrainte informatique imposée par SIGAL peut être insuffisamment contraignante par rapport à la contrainte attendue. Elles sont représentées par des astérisques sur la figure C. Le choix des contraintes attendues pour ces relations est à discuter en fonction du dispositif étudié.

Le tableau A permet d'explicitier les relations pour lesquelles il convient de discuter de l'intérêt de choisir des contraintes relationnelles différentes (et plus strictes) que celles imposées. Ce tableau précise également la nature de la contrainte imposée par SIGAL.

Tableau A : Contraintes informatiques imposées par SIGAL selon les types de relation du SRD

Type de relation	Contraintes informatiques imposées	
	Type	Commentaires
1 : Nombre d'interventions par EES	0,n	Absence de contraintes imposées par SIGAL.
2 : Nombre d'EES par intervention	1,n	Il y a au moins une EES par intervention, et il peut y en avoir plusieurs
3 : Nombre d'EES par RAI	1,n	Il y a au moins une EES par intervention, et il peut y en avoir plusieurs
4 : Nombre de RAI par intervention	0,n	Absence de contraintes imposées par SIGAL. L'intervention peut être saisie manuellement (sans RAI) ou à partir d'un ou de plusieurs RAI.
5 : Nombre de RAI par EES	0,n	Absence de contraintes imposées par SIGAL.

La nécessité de discuter des règles de décision permettant de fixer les contraintes relationnelles listées au tableau A se justifie par les différents points suivants :

- Concernant la **relation 1**, qui détermine le nombre d'interventions par EES : dans tous les cas, une intervention doit être créée pour chaque entité, et s'il y en a plusieurs, ce sont des doublons. Ainsi, il est nécessaire d'imposer une relation attendue de type 1,1.
- Concernant les **relations 2 et 3**, qui déterminent le nombre d'EES par intervention et par RAI : il est nécessaire de savoir si certains commémoratifs sont à renseigner au niveau de l'EES. En effet, SIGAL ne permet de renseigner les commémoratifs qu'au niveau de l'intervention et non au niveau du RAI ou de l'animal (ou du groupe d'animaux). Pour une intervention, il n'y a donc qu'une donnée par type de commémoratifs. Or, certains commémoratifs donnent des informations à l'échelle de l'animal (ou du groupe d'animaux) alors que plusieurs animaux (ou groupes d'animaux) peuvent être rassemblés dans la même intervention. Par exemple, si un VS fait une visite pour deux avortements ayant eu lieu dans la même exploitation, ces deux avortements peuvent être renseignés dans la même intervention, mais alors une seule date d'avortement et une seule durée de gestation pourront être renseignées.
De ce fait, s'il existe des commémoratifs au niveau de l'EES, il est nécessaire de n'avoir qu'une EES par intervention et qu'une EES par RAI.
Dans le cas où il n'y a pas de commémoratifs au niveau de l'EES, il faut se demander s'il est nécessaire de n'avoir qu'un EES par intervention et par RAI.
- Concernant la **relation 4**, qui détermine le nombre de RAI par intervention : on peut demander à avoir au moins un RAI par intervention. Ceci impose dès lors la transmission des résultats sous forme informatique et met donc de côté les interventions saisies entièrement par la DDecPP ou le GDS. Cependant, ce mode de fonctionnement est peu présent aujourd'hui.
Pour le nombre maximum de RAI par intervention, deux cas sont à envisager. S'il y a des commémoratifs au niveau de l'EES, il est nécessaire de n'avoir qu'un RAI par intervention (s'il y a plusieurs RAI, il y a automatiquement plusieurs EES). Dans le cas où il n'y a pas de commémoratifs au niveau de l'EES, il faut se demander si rattacher plusieurs RAI à une intervention entraîne une perte d'information, et à l'inverse, quels sont les avantages de rattacher plusieurs RAI à une même intervention. Par exemple, dans le cadre de la surveillance active de la brucellose, il n'y a qu'une intervention par an. Or, il est possible de réaliser la surveillance en plusieurs fois, donc sur plusieurs visites : il est donc intéressant d'avoir plusieurs RAI (et plusieurs EES) pour cette intervention.
- Concernant la **relation 5**, c'est-à-dire le nombre de RAI par EES : lorsque l'Etat prend en charge les prélèvements réalisés par le vétérinaire et que plusieurs analyses sont faites sur le même prélèvement, imposer un unique RAI par EES évite les doubles saisies (pour le laboratoire) et limite les risques de double paiement du prélèvement.

Les relations présentées dans le SRD pour lesquelles le type de relation souhaité est plus contraint que ce qui est prévu informatiquement (donc celles pour lesquelles des écarts possibles au SRD cible peuvent être observés) doivent être listées.

→ **Compléter la sous-section 1-D – Réalisation d'un schéma relationnel des données**

Exemple d'évaluation : le dispositif de déclaration des avortements

Dans le cas du dispositif de DA, il y a des commémoratifs à l'échelle du bovin : la date d'avortement et la durée de gestation. Les contraintes retenues sont présentées dans le tableau B.

Tableau B : Contraintes souhaitées pour les différentes relations du SRD pour le dispositif de DA

Type de relation	Contraintes souhaitées
1 : Nombre de DA par intervention / par RAI	1,1
2 : Nombre de RAI par intervention	1,1
3 : Nombre d'interventions par DA	1,1
4 : Nombre de RAI par DA	1,1

1.3 Sélection et description des données (section 2 du formulaire)

Cette étape concerne plus particulièrement les données collectées dans le cadre du dispositif de surveillance évalué. Il s'agit de les sélectionner, puis de les décrire et enfin de les catégoriser.

1.3.1 Sélection des données

Il est tout d'abord nécessaire de faire un inventaire de toutes les données disponibles dans SIGAL (nom et définition des données collectées dans le cadre du dispositif de surveillance étudié). Cet inventaire doit permettre de sélectionner les données à retenir pour conduire l'évaluation, c'est-à-dire :

- les données collectées à des fins de surveillance et non de gestion : elles font l'objet de l'évaluation de la qualité ;
- les données qui pourraient être nécessaires à la réalisation de l'évaluation mais dont la qualité n'est pas étudiée (comme par exemple l'identifiant du département, généré automatiquement par SIGAL et qui permet de mener l'évaluation à l'échelle départementale).

Certaines données sont à étudier dans tous les cas : les différents commémoratifs et les données relatives aux analyses de laboratoire (prélèvement, matrice, analyte, méthode, résultat, interprétation). Il faut y ajouter l'identifiant de l'animal ou du groupe d'animaux et l'identifiant de la structure rattachée à l'intervention, bien que pour les IP ces données soient enregistrées par avance dans la DAI, générée automatiquement par SIGAL. En effet, du fait de la possibilité de délais plus ou moins importants entre l'envoi de la DAI et la réalisation de l'intervention (jusqu'à un an) ainsi que de la possibilité dans certains cas de rajouter des unités épidémiologiques⁷, l'évaluation de la qualité de ces données reste nécessaire pour les IP.

A ces données étudiées obligatoirement peuvent s'en ajouter d'autres, en fonction des données collectées et saisies pour le dispositif étudié.

⁷ Par exemple, si un bovin a été introduit dans l'exploitation mais n'apparaît pas sur le DAP (pour une raison de mise à jour du DAP), il est possible de réaliser un prélèvement sur cet animal sous réserve de l'ajouter manuellement sur le DAP.

Une fois les données nécessaires à l'évaluation sélectionnées, les critères d'extraction peuvent être précisés : ils doivent inclure la nature de l'extraction et la période d'étude.

Exemple d'évaluation : le dispositif de déclaration des avortements

L'évaluation de la qualité des données de DA chez les bovins a porté sur les données collectées à des fins de surveillance, c'est à dire :

- la date d'intervention
- le numéro d'ordre
- l'identifiant de l'exploitation – atelier
- l'identifiant bovin
- la date d'avortement
- la durée de gestation
- la matrice prélevée
- le motif de non analyse
- l'état complet du plan d'analyse
- la matrice analysée
- l'analyte
- la méthode
- l'analyse de confirmation
- le résultat
- la date de validation des résultats
- la date du RAI
- l'état complet du RAI
- le laboratoire et le laboratoire sous-traitant
- l'interprétation

Ainsi, la qualité des données suivantes n'a pas été étudiée :

- les numéros de lot et d'échantillon, qui sont des données internes au laboratoire ;
- la catégorie fiscale du véhicule, le nombre de km parcourus par le VS pour réaliser sa visite et le motif de non-paiement, ces données n'étant utiles que pour la gestion locale par les DDecPP (et le paiement du VS) ;
- le numéro de Lims, l'identifiant du RAI, le numéro d'intervention et le département de saisie, qui sont des données générées automatiquement par les systèmes informatiques ;
- le type d'intervention (« enregistrement d'une déclaration d'avortement bovin ») et le plan d'analyse utilisé (« plan EBRUBVA »), qui sont les données à partir desquelles a été faite l'extraction des données de SIGAL ;
- la date de création de l'intervention, générée automatiquement, qui est une donnée disponible dans SIGAL mais qui ne fait pas partie de l'extraction utilisée pour l'évaluation de la qualité des données ;
- la synthèse et la décision, qui ne sont utiles que pour la gestion locale et qui ne sont accessibles qu'aux DDecPP ;
- la qualité du numéro d'intervention et du département de saisie, sachant que ces données ont été utilisées pour évaluer la qualité d'autres données.

1.3.2 Description et catégorisation des données

Les caractéristiques informatiques des données sélectionnées doivent être décrites précisément par l'intermédiaire du renseignement de métadonnées⁸ : nom, définition, obligation, format, valeurs possibles, mise à jour, source de la donnée. Les significations de ces métadonnées sont précisées dans le tableau C. La description des données pourra éventuellement être validée avec les acteurs du dispositif.

Tableau C : Définitions des métadonnées

Métadonnée	Définition
Nom de la donnée	Caractérise la dénomination retenue dans le jeu de données étudié
Définition	Définit précisément la donnée
Obligation	Précise si le renseignement de la donnée est imposé informatiquement, en fonction de la source
Format	Précise si le format de la donnée est contraint informatiquement
Valeurs possibles	Précise si la valeur doit faire partie d'une liste de valeurs possibles (liste positive, ou intervalle de valeurs)
Mise à jour	Précise si une mise à jour de la donnée est possible (ou sa correction a posteriori en cas d'erreur par exemple), et si oui par qui
Source de la donnée	Situe la source de la donnée analysée (donnée renseignée directement dans SIGAL à partir de l'intervention ou dans le RAI par le laboratoire)

Cette description à l'aide des métadonnées permet de définir les degrés de structuration et de renseignement de chaque donnée afin d'aider au choix des dimensions qui leur sont appliquées (tab D). En effet, certaines dimensions n'ont pas d'intérêt d'être étudiées pour certaines données (par exemple, la complétude d'une donnée dont le renseignement est rendu obligatoire par des contraintes informatiques est obligatoirement de 100%).

Tableau D : Définitions des critères de tri des données

Niveau de renseignement	Valeur manquante possible		Valeur manquante impossible
Degré de structuration [6]	Données structurées : le format et la valeur à renseigner sont prédéfinis (listes déroulantes)	Données semi-structurées : la structure a un certain degré de flexibilité (par exemple une donnée peut être obligatoirement sous un format date et avec une valeur libre)	Données non structurées : la valeur est entièrement libre

Les mêmes données peuvent être enregistrées soit par le laboratoire (à partir de leur Lims et exportées dans SIGAL via un RAI) soit par la DDecPP ou le GDS (directement dans SIGAL). Les contraintes informatiques peuvent diverger entre les Lims et SIGAL, conduisant ainsi à un degré de structuration et un niveau de renseignement des données variables. Dans ce cas, il convient de retenir le degré de structuration et le niveau de renseignement les moins contraignants.

⁸ Une métadonnée est une donnée servant à décrire ou définir une autre donnée.

→ Compléter la section 2 – Sélection et description des données

Exemple d'évaluation : le dispositif de déclaration des avortements

La description des données collectées dans le cadre du dispositif de DA est présentée en annexe III. Leur catégorisation est présentée tableau E.

Dans le cas du dispositif de DA, la date d'avortement et la durée de gestation sont soumises à des contraintes informatiques dans SIGAL (données semi-structurées) alors qu'aucune contrainte informatique ne s'applique dans les Lims des laboratoires (données non structurées). Cette absence de contrainte persiste au moment du rapatriement de ces données dans SIGAL pour créer l'intervention. En conséquence, la date d'avortement et la durée de gestation ont été traitées comme des données non structurées.

Tableau E : Catégorisation des données de DA selon leur degré de structuration et leur niveau de renseignement

	Caractéristiques des données	
	Structuration	Renseignement
Date du RAI	Structurées	Obligatoire
Etat complet du RAI		
Etat complet du plan d'analyse		
Numéro du VS		Non obligatoire (mais la présence de l'un oblige celle des autres)
Nature du prélèvement		
Matrice analysée		
Analyte		
Méthode		
Résultat		
Analyse de confirmation		
Laboratoire		Obligatoire
Identifiant de l'exploitation		
Date de validation des résultats	Semi-structurées	
Date de visite		
Identifiant du bovin	Non structurées	Non obligatoire
Atelier	Structurées	
Motif de non analyse		
Interprétation des résultats		
Laboratoire sous-traitant		
Date d'avortement	Non structurées	
Durée de gestation		

EVALUATION QUALITATIVE DE LA QUALITE DES DONNEES (SECTION 3 DU FORMULAIRE)

Cette section représente le volet qualitatif de l'évaluation de la qualité des données. Cette évaluation doit permettre de valider les descriptions du dispositif et du circuit des données puis les acteurs impliqués dans le dispositif. Par ailleurs, elle permettra d'aider à l'interprétation des résultats

quantitatifs, notamment via la mise en évidence de facteurs influençant la qualité des données, et d'ouvrir des pistes d'amélioration.

Les dimensions retenues pour cette étude qualitative portent sur la compréhension, la pertinence, la crédibilité et l'accès / utilisation des données collectées.

Cette évaluation se fait au travers d'entretiens avec tous les types (ou familles) d'acteurs du dispositif, identifiés précédemment (section 1, liste et schéma de circulation des données). De manière indicative, ces entretiens durent entre 30 minutes et 1h30 en fonction du type d'acteurs et du nombre de données à évaluer. Afin d'appréhender la diversité des situations et des points de vue (notamment concernant les règles de gestion et les types de relation entre acteurs), il convient de mener cette étude dans plusieurs départements.

Exemple d'évaluation : le dispositif de déclaration des avortements

En ce qui concerne la DA, l'évaluation des dimensions de second niveau a été réalisée lors de déplacements dans trois départements, le Rhône, le Puy-de-Dôme et la Haute-Savoie. Ces départements ont été choisis pour des raisons pratiques (proximité géographique), pour leur diversité (présence d'une délégation au GDS dans le Rhône et pas dans les deux autres départements) et en raison de l'accord des acteurs de terrain à participer à cette étude. Elle s'est basée sur des entretiens avec les différents acteurs prenant part à la collecte et/ou à la saisie des données : huit VS, quatre personnes travaillant dans un laboratoire, deux dans un GDS et quatre en DDecPP. Les entretiens se sont déroulés en deux parties. La première partie a été consacrée à la présentation de la définition de la qualité des données et des dimensions étudiées. Ensuite, pour chaque donnée, une note entre 0/3 et 3/3 a été attribuée par dimension de second niveau. Cette notation a été associée à une discussion.

L'entretien peut se dérouler suivant les étapes suivantes :

1. Présentation des objectifs de l'évaluation et de son intérêt

Un document écrit de présentation de la démarche est souhaitable, et il est laissé à la personne rencontrée. En plus d'informations ciblées sur l'évaluation menée (objectifs et intérêts), il doit expliquer la notion de qualité des données et son intérêt plus global.

Un document type est présenté en première partie de l'annexe IV.

2. Discussion et validation des étapes précédentes

Le schéma de circulation des données et la description des données à l'aide de métadonnées réalisés précédemment peuvent être validés et éventuellement corrigés. Chaque acteur ne peut être interrogé que sur les données qu'il connaît.

3. Données non saisies

La collecte de certaines données peut être demandée sans toutefois que leur enregistrement soit prévu informatiquement, par exemple l'âge du bovin. Dans ce cas, l'intérêt de la collecte de ces données doit être discuté. Il s'agit de déterminer si ces données peuvent être utiles, et plus précisément si elles méritent ou non la mise en place d'une saisie. En outre, il convient de

s'interroger sur l'intérêt de collecter et saisir des données qui ne sont actuellement pas collectées (et donc pas saisies).

Par la suite, seules les données collectées et saisies sont prises en compte.

4. Compréhension des données

La personne menant l'entretien doit évaluer la compréhension des données collectées, saisies ou accessibles par la personne interrogée via SIGAL ou une autre interface.

Il s'agit donc dans un premier temps d'identifier ces données puis d'attribuer une note de compréhension pour chaque donnée concernée. Pour cela, la personne menant l'entretien demande à la personne interrogée de définir la donnée (définition au sens strict et format attendu), puis attribue une note à la donnée étudiée de la manière suivante :

- la définition et le format attendu sont tout à fait ceux définis préalablement par l'équipe d'évaluation : 3/3 ;
- la définition et le format attendu sont proches de ceux définis préalablement par l'utilisateur, cela n'entraîne pas de gros problème de sens : 2/3 ;
- la donnée et le format attendu ne sont qu'en partie compris : 1/3 ;
- la donnée et le format attendu ne sont pas compris : 0/3.

Pour les données pour lesquelles la note de compréhension n'est pas de 3/3, la définition doit être reprécisée avant de poursuivre afin de s'accorder sur le sens de la donnée.

5. Pertinence, crédibilité et accès / utilisation des données

Ces dimensions sont évaluées directement par la personne interrogée. La pertinence et l'accès / utilisation ne sont évalués que si la personne interrogée a accès à la donnée ; la crédibilité ne l'est que si la personne interrogée collecte ou saisit la donnée.

Un document type de support pour cette étape est présenté en seconde partie de l'annexe IV.

De la même manière que pour la compréhension, l'évaluation se fait sous la forme d'une notation. Les règles de notation sont présentées dans le tableau F.

Tableau F : Règles de notation pour la personne interrogée

Dimension	Question	Note			
		3/3	2/3	1/3	0/3
Pertinence	Pensez-vous que cette donnée soit utile, que ce soit pour la surveillance ou pour vous (gestion, communication...) ?	Oui, elle est indispensable	Plutôt oui	Pas vraiment	Non, pas du tout
Crédibilité	Quelle confiance avez-vous dans l'exactitude (et la précision) de cette donnée ?	Oui, totalement	Oui, assez	Pas vraiment	Non, pas du tout
Accès / utilisation	Avez-vous facilement accès à cette donnée ? Est-ce facile de l'utiliser ?	Très facile	Plutôt facile	Assez compliqué	Très compliqué

La notation doit s'accompagner d'une discussion et d'une argumentation par la personne interrogée, nécessaires pour identifier les éventuels facteurs influençant la qualité des données. En particulier, lors de la notation, il faut recueillir :

- Pour la pertinence : des informations sur l'utilité de la donnée ;
- Pour la crédibilité : si elle est jugée mauvaise, les raisons de ce jugement (manque d'objectivité par exemple) ;
- Pour l'accès / utilisation : est-ce que la donnée est utilisée, et si oui à quelle fréquence et pour quoi.

→ **Compléter la section 3 – Evaluation qualitative de la qualité des données**

EVALUATION QUANTITATIVE DE LA QUALITE DES DONNEES (SECTIONS 4 ET 5 DU FORMULAIRE)

Les dimensions utilisées pour l'étude quantitative portent sur la complétude, le format, la validité, la vitesse de circulation des données et la cohérence.

D'autres dimensions sont trouvées dans la littérature et peuvent éventuellement être étudiées, telles que la volatilité, la sécurité et la précision. Toutefois, il convient de manière générale d'essayer de limiter le nombre de dimensions. En outre, la sécurité est une dimension qui se rapporte davantage au système informatique qu'aux données elles-mêmes ; cette notion est traitée lorsque les droits d'écriture et de lecture sont précisés (partie II.A.1.d)). La précision recoupe dans une certaine mesure le format et/ou la validité. La volatilité semble sans objet pour les données collectées dans le cadre de dispositif de surveillance : pour chaque acte réalisé les données sont définitives, elles n'évoluent pas dans le temps.

1.4 Evaluation des spécifications générales d'enregistrement (section 4 du formulaire)

Cette section correspond à la première étape de l'évaluation quantitative de la qualité des données. Pour cette étape, l'intégralité des enregistrements présents dans l'extraction doit être étudiée.

Il s'agit d'étudier l'enregistrement des données dans SIGAL, afin de sélectionner les entités élémentaires de surveillance enregistrées de façon à permettre une analyse fiable des données qu'elles contiennent. En effet, les erreurs d'enregistrement de certaines données, les absences de contrôles informatiques et la variabilité des modalités de gestion informatique des EES peuvent conduire à des doublons ou bien mettre en doute leur réalisation effective : par exemple, si l'intervention est enregistrée dans une exploitation dont l'activité a cessé (mais qui est restée « ouverte » dans SIGAL), on peut se demander si cette intervention a réellement eu lieu ou s'il s'agit d'une erreur de saisie.

1.4.1 Choix des critères d'évaluation

Plusieurs critères permettent d'évaluer le respect des spécifications générales d'enregistrement, en particulier :

- la validité du numéro d'identification de la structure rattachée à l'intervention, par exemple l'identifiant EDE. Cet identifiant est unique pour l'exploitation ou spécifique de l'atelier si l'exploitation comporte plusieurs ateliers localisés sur différentes communes. La vérification de ce critère nécessite une base de référence (par exemple comparaison avec la BDNI) à laquelle il

faut donc avoir accès. Cette étape n'a pas lieu d'être si la base de référence est SIGAL, ce qui est par exemple le cas des ateliers de volailles ;

- l'absence d'interventions en doublon : pour une EES, il n'y a bien qu'une seule intervention ;
- la cohérence entre la date de la visite et l'envoi du RAI (la date de visite doit précéder la date d'envoi du RAI).

D'autres critères peuvent être ajoutés, notamment le respect des relations pour lesquelles les contraintes souhaitées sont plus fortes que les contraintes informatiques existantes (s'appuyer sur le tableau rempli en section 1-D du formulaire).

1.4.2 Construction et mesure des indicateurs

Pour chaque critère étudié, un indicateur est construit et mesuré au niveau national et au niveau départemental. Les structures des différents indicateurs à utiliser sont décrites dans le tableau G.

Tableau G : Indicateurs relatifs au respect des spécifications générales d'enregistrement

Critère	Indicateur (numérateur / dénominateur)
Validité de l'identifiant de la structure à laquelle est rattachée l'intervention *	Nombre d'interventions qui respectent le critère / Nombre total d'interventions
Absence d'interventions en doublon **	Nombre d'EES qui respectent le critère / Nombre total d'EES
Date de visite antérieure à la date de RAI (ou de tous les RAI s'il y en a plusieurs pour une intervention donnée)	Nombre d'interventions qui respectent le critère / Nombre total d'interventions
Autres relations non contraintes	Nombre de couples de données qui respectent la relation / Nombre total de couples de données

* C'est-à-dire l'exploitation ou l'atelier

** Des interventions sont en doublon s'il y en a plus d'une par EES

Une précision doit être apportée sur les modalités de calcul de l'indicateur relatif aux interventions en doublon. Une EES est définie comme l'association de la date de la visite, du lieu de la visite et de la ou des unités épidémiologiques concernées par l'acte. Lorsqu'un seul animal ou groupe d'animaux est inclus dans l'EES (par exemple une déclaration d'avortement ne concerne qu'un animal), la recherche de doublons s'effectue effectivement à partir de ce triplet. Cependant, dans le cas où l'EES peut inclure plusieurs unités épidémiologiques, les doublons peuvent être recherchés pour des raisons pratiques uniquement sur le couple date et lieu de visite. C'est par exemple le cas des prophylaxies : on peut se contenter de vérifier qu'il n'y a bien qu'une visite au même endroit le même jour, sans pour autant s'assurer que chaque animal ou chaque groupe d'animaux n'est bien prélevé qu'une fois.

→ Compléter la section 4 – Evaluation des spécifications générales d'enregistrement

Exemple d'évaluation : le dispositif de déclaration des avortements
Les DA (qui sont ici les EES) respectant les spécifications générales d'enregistrement sont celles pour lesquelles :

- le schéma relationnel des données (SRD) est respecté, c'est-à-dire que :
 - o le RAI est rattaché à une intervention (et une seule) ;
 - o lorsque le RAI est utilisé pour créer l'intervention, un seul bovin est attribué à un seul RAI, lui-même attribué à une seule intervention : les interventions et RAI correspondant à deux bovins (deux identifiants différents ou nuls) ont été supprimées pour la deuxième partie de l'évaluation quantitative ;
 - o lorsque l'intervention est directement créée dans SIGAL, un seul bovin est attribué à une seule intervention : les interventions correspondant à deux bovins avec deux identifiants différents (ou nuls) ont été supprimées ;
 - o une seule intervention est rattachée à une DA ;
- les interventions sont rattachées à un numéro d'EDE valide ; ainsi, lorsque l'établissement de rattachement était identifié à partir d'un autre identifiant qu'un EDE (SIRET ou SIREN par exemple), l'intervention a été supprimée car ne permettant pas de s'assurer de la présence effective de bovins dans cet établissement ;
- la date d'intervention est antérieure à la date de RAI : la visite lors de laquelle les prélèvements sont réalisés a obligatoirement lieu avant l'exportation des résultats.

Ainsi, afin de vérifier que les DA présentes dans l'extraction disponible respectent les spécifications générales d'enregistrement, plusieurs relations du SRD non contraintes informatiquement ont été vérifiées : l'unicité des liens entre une intervention et un RAI, entre un bovin et une intervention, et enfin entre un avortement et une intervention. Cette analyse devait permettre d'évaluer les modalités de saisie dans SIGAL et le niveau de validation des données à l'échelle de la DA. De plus, la cohérence entre la date de visite et la date de RAI a été étudiée.

Chaque critère de sélection correspond à un indicateur (Tableau H).

Tableau H : Indicateurs retenus pour évaluer le respect du SRD

Critères	Indicateur
Un bovin par intervention	Proportion (en %) d'interventions concernant un seul bovin
Une intervention par avortement	Proportion (en %) d'avortements pour lesquels il y a une seule intervention
Un RAI par intervention	Proportion (en %) d'interventions auxquelles est rattaché un seul RAI
Cohérence date de visite – date de RAI	Proportion (en %) d'interventions pour lesquelles la date de visite est antérieure à la date de RAI

1.5 Evaluation quantitative de la qualité des données respectant les spécifications générales d'enregistrement (section 5 du formulaire)

C'est la seconde étape de l'évaluation quantitative. Elle doit être réalisée sur un jeu de données nettoyé, c'est-à-dire ne comprenant que les EES respectant l'intégralité des spécifications générales d'enregistrement. Ainsi, par la suite, le nombre total d'EES correspond au nombre total d'entités respectant les spécifications générales d'enregistrement.

Les dimensions à étudier sont la vitesse de circulation des données, la complétude, le format, la validité et la cohérence (cf. glossaire).

L'évaluation est conduite à deux niveaux, à l'échelle des données et à l'échelle des entités élémentaires de surveillance. La démarche est progressive (fig D).

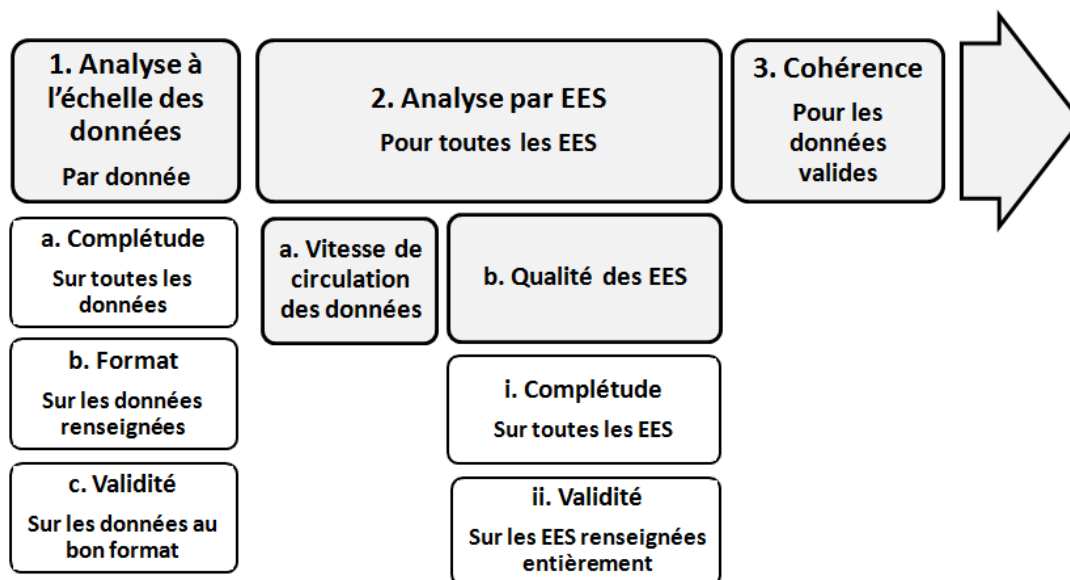


Figure D : Schéma de la démarche d'évaluation de la qualité des données pour les EES respectant les spécifications générales d'enregistrement

1.5.1 A l'échelle des données (sous-section 5-A du formulaire)

Cette étape correspond à l'évaluation de la complétude, du format et de la validité des données. Pour cela, il est nécessaire de sélectionner les couples « donnée / dimension » pour lesquels des indicateurs vont être construits et mesurés. La démarche doit être progressive (fig D).

Ce choix dépend des caractéristiques des données, définies dans la partie A.3.2. En effet, la complétude n'a d'intérêt à être étudiée que pour les données pour lesquelles le renseignement est non obligatoire, le format que pour les données non structurées et la validité que pour les données pour lesquelles une liste de valeurs possibles est disponible. Les autres couples sont sans objet. Par ailleurs, ce choix s'appuie sur les objectifs de l'évaluation et ce qui est attendu de l'analyse des résultats de l'évaluation.

Le détail de la construction des indicateurs est précisé dans le tableau I.

Tableau I : Indicateurs relatifs à l'évaluation à l'échelle des données

Dimensions étudiées	Données concernées par l'évaluation	Indicateurs (numérateur / dénominateur)
Complétude	Uniquement des données pour lesquelles le renseignement n'est pas obligatoire	Nombre de données renseignées / nombre de données dont le renseignement est attendu
Format	Uniquement des données pour lesquelles le format est non contraint (définir le format attendu pour chacune d'entre elle)	Nombre de données au bon format / nombre de données renseignées
Validité	Uniquement des données pour lesquelles le	Nombre de données valides /

	format est complètement contraint (définir ce qu'est une donnée valide)	nombre de données renseignées et au bon format
--	---	--

Le nombre de données dont le renseignement est attendu peut être différent en fonction du type de données. Par exemple, le nombre de « durées de gestation » à renseigner, associé au nombre de DA, peut être différent du nombre d'analyses à renseigner : une DA peut faire l'objet de plusieurs analyses.

→ **Compléter la sous-section 5-A – Evaluation quantitative à l'échelle des données**

Exemple d'évaluation : le dispositif de déclaration des avortements

Les données pour lesquelles la complétude, le format et la validité ont été étudiés sont détaillées dans le tableau J.

La complétude a été étudiée pour l'ensemble des données pour lesquelles le renseignement n'est pas obligatoire, à l'exception de l'atelier, le motif de non analyse et le laboratoire sous-traitant. En effet, le renseignement de ces données peut être sans objet et donc leur complétude n'a pas de signification. Par exemple, si l'analyse a été réalisée, il n'y a logiquement pas de motif de non analyse. De la même façon, le respect du format n'a été étudié que pour les données non structurées. Enfin, la validité n'a pu être étudiée que pour les données pour lesquelles une liste de valeurs possibles était disponible. En effet, cette dimension correspond à la concordance entre la donnée et un ensemble de valeurs considérées comme « valides ».

Tableau J : Données pour lesquelles la complétude, le format et la validité ont été retenus.
 NB : Les cases grisées représentent les couples « donnée / dimension » dont la qualité a été étudiée.

	Caractéristiques des données		Dimensions étudiées		
	Structuration	Renseignement	Complétude	Format	Validité
Date du RAI	Structurées	Obligatoire			
Etat complet du RAI					
Etat complet du plan d'analyse					
Numéro du VS					
Nature du prélèvement		Non obligatoire (mais si l'une des donnée est renseignée, les autres doivent l'être)			
Matrice analysée					
Analyte					
Méthode					
Résultat					
Analyse de confirmation		Obligatoire			
Laboratoire					
Identifiant de l'exploitation					
Date de validation des résultats					
Date de visite	Semi-structurées				
Identifiant du bovin	Non structurées	Non obligatoire			
Atelier	Structurées				
Motif de non analyse					
Interprétation des résultats					
Laboratoire sous-traitant					
Date d'avortement	Non structurées				
Durée de gestation					

Le choix des couples « donnée / dimension » a également été fonction des objectifs visés, c'est à dire :

- pour l'identifiant bovin : s'assurer d'un niveau de traçabilité élevé, indispensable en cas de confirmation d'une suspicion ;
- pour la date d'avortement : s'assurer d'une fiabilité suffisante du délai d'obtention des résultats et donc de la réactivité du dispositif ;
- pour la durée de gestation : permettre une analyse du type d'avortement surveillé à mettre en rapport avec la clinique de la brucellose ;
- pour la nature du prélèvement : permettre une vérification de la bonne application de la note de service définissant les prélèvements autorisés (une prise de sang et éventuellement un écouvillon endo-cervical) (Anonyme, 2010b ; Anonyme, 2010a) ;
- pour les données relatives aux analyses de laboratoire : s'assurer de la validité des analyses réalisées et donc de la possibilité de les interpréter à l'échelle de l'animal (indirectement via la vérification du respect de la fiche de plan) ;

- pour l'interprétation du laboratoire : permettre une analyse plus efficace et efficiente du dispositif, en s'intéressant non plus aux résultats d'analyses mais à l'interprétation.

Les indicateurs correspondants sont présentés dans le Tableau K. L'unité utilisée est la DA ou l'analyse selon que le support de la donnée étudiée est le RAI ou l'intervention.

Tableau K : Indicateurs relatifs aux dimensions de niveau 1 retenus pour l'évaluation de la qualité des données à l'échelle des données.

Dimension (et donnée concernée)	Indicateur (numérateur / dénominateur)	Jeu de données utilisé
Complétude (date d'avortement, durée de gestation)	Nombre de DA avec la donnée renseignée / Nombre total de DA	Jeu de données nettoyé.
Complétude (identifiant bovin, nature du prélèvement, interprétation des résultats)	Nombre d'analyses avec la donnée renseignée / Nombre total d'analyses	
Format (identifiant bovin, date d'avortement, durée de gestation)	Nombre de DA avec la donnée au bon format / Nombre total de DA	Jeu de données nettoyé, et DA pour lesquelles la donnée étudiée était renseignée.
Validité (durée de gestation et identifiant bovin)	Nombre de DA avec une donnée valide / Nombre total de DA	Jeu de données nettoyé, et les DA ou les analyses pour lesquelles la donnée étudiée était renseignée et au bon format.
Validité (nature du prélèvement, matrice analysée, méthode, résultat)	Nombre d'analyse avec une donnée valide / Nombre total d'analyse	

1.5.2 A l'échelle de l'entité élémentaire de surveillance (sous-section 5-B du formulaire)

Vitesse de circulation des données

La première étape de l'évaluation à cette échelle est l'étude de la vitesse de circulation des données. Pour cette dimension, les données sont étudiées à partir de l'étude de différents délais, et en particulier les deux délais suivants :

- Le délai d'obtention des résultats, qui correspond au délai entre :
 - o pour les INP : la date de l'événement et la date de validation des résultats,
 - o pour les IP : la date de réalisation de la visite et la date de validation des résultats ;
- Le délai de disponibilité des données : délai entre la date de validation des résultats et la date d'exportation du RAI.

Ces deux délais peuvent éventuellement être adaptés en fonction des dates disponibles. Notamment, pour les INP, la date d'exportation du RAI peut être remplacée par la date de création de l'intervention.

Il est conseillé de ne conserver que ces deux délais (tels que présentés ou après adaptation) ; il faut garder à l'esprit qu'il n'est pas nécessaire ni souhaitable de multiplier les indicateurs et de calculer tous les délais possibles.

Les indicateurs d'évaluation de la vitesse de circulation des données correspondent à la moyenne de chacun des délais, exprimée en jours et calculée aux niveaux national et départemental.

Exemple d'évaluation : le dispositif de déclaration des avortements	
<p>Pour évaluer la vitesse de circulation des données, deux délais ont été utilisés : le délai entre la date d'avortement et la date de validation des résultats, puis le délai entre la date de validation des résultats et la date d'envoi du RAI dans SIGAL. Les indicateurs utilisés étaient les moyennes des délais correspondants (Tableau L).</p> <p>Cette analyse devait permettre de quantifier dans quelle mesure le dispositif permettrait réellement de détecter précocement un nouveau foyer (réactivité du dispositif), et de connaître les délais de mise à disposition des données dans SIGAL, ce qui peut influencer par exemple la qualité de la validation des données (plus une donnée est ancienne, moins sa validation est aisée).</p> <p>Tableau L : Indicateurs retenus pour évaluer la vitesse de circulation des données</p>	
Dimension	Indicateur
Délai d'obtention des résultats	Délai moyen entre la date d'avortement et la date de validation des résultats (en jours)
Disponibilité des données	Délai moyen entre la date de validation des résultats et la date d'envoi du dernier RAI de l'intervention (en jours)

Complétude et validité

Dans un second temps, il s'agit d'estimer la proportion d'entités entièrement de bonne ou de mauvaise qualité du point de vue de deux dimensions, la complétude et la validité. La démarche à suivre est progressive : la complétude est d'abord étudiée, puis la validité est étudiée uniquement sur les entités élémentaires de surveillance entièrement renseignées. Les indicateurs correspondants sont présentés dans le tableau M.

Tableau M : Indicateurs relatifs à l'évaluation de la complétude et de la validité à l'échelle des entités élémentaires de surveillance.

Dimension	Indicateur (numérateur / dénominateur)
Complétude	Nombre d'entités pour lesquelles toutes les données sont renseignées / nombre total d'entités
Validité	Nombre d'entités pour lesquelles toutes les données sont valides / nombre d'entités où elles sont toutes renseignées

Lorsqu'une entité a plusieurs données de même type (par exemple, deux analyses), il est nécessaire de vérifier que toutes les données rattachées à l'entité répondent au critère.

→ Compléter la sous-section 5-B – Evaluation à l'échelle de l'entité élémentaire de surveillance

Exemple d'évaluation : le dispositif de déclaration des avortements

Le tableau N rassemble les indicateurs qui ont été retenus pour l'évaluation de la qualité des données à l'échelle des DA.

Tableau N : Indicateurs relatifs aux dimensions de niveau 1 retenus pour l'évaluation de la qualité des données à l'échelle des DA.

Dimension	Indicateur
Complétude	Proportion de DA pour lesquelles toutes les données étudiées sont renseignées
Validité	Proportion de DA complètes pour lesquelles toutes les données étudiées sont valides

1.5.3 Etude de la cohérence (sous-section 5-C du formulaire)

L'étude de cette dimension doit permettre de s'assurer que les données d'une entité élémentaire de surveillance sont cohérentes entre elles. La cohérence n'est étudiée que sur les données renseignées, au bon format et valides. Elle peut être conduite à différents niveaux selon les données concernées (tab O).

Tableau O : Cohérence entre les données à étudier et niveau auquel l'évaluation est conduite

Cohérence	Niveau auquel l'évaluation est conduite	IP	INP
1. Nature du prélèvement - matrice analysée	Analyse	X	X
2. Matrice analysée - analyte recherché - méthode d'analyse (MAM = une analyse)	Analyse	X	X
3. MAM - motif de non analyse	Analyse	X	X
4. MAM - analyse de confirmation	Analyse	X	X
5. Différentes analyses d'un RAI	RAI	X	X
6. Résultats – interprétation des résultats	RAI	X	X
7. Date de l'évènement – date de visite	Entité élémentaire de surveillance		X

Les croix indiquent si la cohérence doit être étudiée ou non.

La démarche est progressive : l'étude des cohérences 3, 4 et 5 ne se fait que pour les analyses pour lesquelles la cohérence 2 est effective, et l'étude de la cohérence 6 ne se fait que pour les entités pour lesquelles la cohérence 5 est effective.

D'autres cohérences peuvent être étudiées, en gardant à l'esprit les objectifs de l'étude (mise en évidence de facteurs influençant la qualité des données) et l'intérêt de limiter le nombre d'indicateurs.

Les indicateurs sont calculés comme le nombre de couples de données respectant la cohérence étudiée rapporté au nombre de couples de données dont la cohérence a été étudiée.

La mesure des indicateurs relatifs aux cohérences 5 et 6 (différents MAM d'un RAI et résultats - interprétation) nécessite une mise en forme du jeu de données. En effet, ces cohérences sont étudiées au niveau du RAI, qui peut comporter plusieurs analyses. Il est donc nécessaire de construire à partir de l'extraction des données, pour laquelle une ligne correspond à un triplet « RAI /

MAM / résultat » (tab P), un tableau dans lequel une ligne correspond à un RAI, et qui peut alors comporter plusieurs MAM (tab Q).

Tableau P : Structuration de la table des données extraite de SIGAL

Nom du RAI	MAM	Résultat	Interprétation
RAI A	MAM A1	Rés A1	Int A
RAI A	MAM A2	Rés A2	Int A
RAI A	MAM A3	Rés A3	Int A
RAI B	MAM B1	Rés B1	Int B
RAI B	MAM B2	Rés B2	Int B

Tableau Q : Structuration des données nécessaire pour mesurer les cohérences 5 et 6

Nom du RAI	MAM 1	Rés 1	MAM 2	Rés 2	MAM 3	Rés 3	Interprétation
RAI A	MAM A1	Rés A1	MAM A2	Rés A2	MAM A3	Rés A3	Int A
RAI B	MAM B1	Rés B1	MAM B2	Rés B2	-	-	Int B

Les colonnes grisées correspondent aux données pour lesquelles l'unité est le RAI, les colonnes blanches correspondent à celles pour lesquelles l'unité est l'analyse.

→ **Compléter la sous-section 5-C – Cohérence entre données**

Exemple d'évaluation : le dispositif de déclaration des avortements	
	<p>En ce qui concerne les DA, les cohérences étudiées ont été les suivantes : la cohérence entre la date d'avortement et la date de visite, la cohérence entre le prélèvement et la matrice analysée, la cohérence entre la matrice analysée, l'analyte recherché et la méthode d'analyse (MAM), la cohérence entre le MAM et le motif de non analyse, la cohérence entre le MAM et l'analyse de confirmation, la cohérence entre les différents MAM d'un même RAI, la cohérence entre les résultats et leur interprétation.</p>

1.6 Construction et mesure d'indicateurs synthétiques (section 6 du formulaire)

Dans les sections précédentes (*à définir après avoir fixé le plan de la procédure*), différents indicateurs spécifiques ont été mesurés aux niveaux national et départemental. Deux indicateurs synthétiques (IS) peuvent être élaborés à partir de ces indicateurs spécifiques, afin d'avoir une vision plus globale de la qualité des données.

Le premier IS permet d'avoir une vision globale du respect des spécifications générales d'enregistrement ; il est calculé à partir des indicateurs relatifs au respect de ces spécifications. Le second IS permet d'avoir une vision globale de la qualité des données et des cohérences entre elles ; il est calculé à partir des indicateurs relatifs à l'évaluation à l'échelle des données, à la vitesse de circulation des données et à la cohérence. Les indicateurs relatifs à l'évaluation de la qualité des données à l'échelle de l'entité élémentaire de surveillance (complétude et validité) ne sont pas inclus

dans le calcul d'IS car ils sont directement dépendants des autres indicateurs spécifiques et n'apporteraient pas d'informations supplémentaires.

1.6.1 Principe de construction des indicateurs

Les indicateurs synthétiques sont mesurés au niveau départemental. Pour chaque indicateur spécifique, une valeur cible doit être définie et à partir de là, un IS est calculé comme le nombre d'indicateurs spécifiques atteignant ou dépassant la valeur cible correspondante rapporté au nombre d'indicateurs spécifiques mesurés :

$$IS = \frac{\text{nombre d'indicateurs atteignant ou dépassant la valeur cible}}{\text{nombre d'indicateurs mesurés}}$$

Ainsi, pour un département, l'IS 1 correspond au nombre d'indicateurs spécifiques relatifs au respect des spécifications générales d'enregistrement atteignant ou dépassant leur valeur cible sur le nombre d'indicateurs spécifiques calculés et relatifs au respect des spécifications générales d'enregistrement. L'IS 2 correspond au nombre d'indicateurs spécifiques relatifs à l'évaluation à l'échelle des données et à ceux relatifs aux cohérences atteignant ou dépassant leur valeur cible sur le nombre d'indicateurs spécifiques calculés et relatifs à l'évaluation à l'échelle des données et à ceux relatifs aux cohérences. Du fait de la démarche progressive, le nombre d'indicateurs mesurés peut être différent entre les départements. Dans ce cas, les IS ont des dénominateurs différents : ils doivent tous être rapportés au même dénominateur pour pouvoir être comparés.

1.6.2 Fixation des valeurs cibles

Une valeur cible pour un indicateur spécifique donné représente la valeur de l'indicateur à partir de laquelle le critère de qualité étudié est considéré bon. Elles sont à adapter en fonction :

- des consignes existantes (formelles ou non) ;
- des objectifs du dispositif ;
- des contraintes de terrain (par exemple, ne pas demander une complétude à 100% si la donnée peut ne pas être connue) ;
- de l'importance de la qualité de la donnée (par exemple la complétude de l'identifiant bovin est plus importante que la complétude de la date d'avortement) ;
- du niveau de prise en compte du facteur humain : des erreurs de saisie et de collecte doivent être tolérées dans une certaine mesure (de l'ordre de 2 à 3% semble un optimum).

Par ailleurs, les valeurs cibles peuvent aussi être fixées en fonction des résultats obtenus pour les indicateurs spécifiques.

Des propositions et des recommandations pour la fixation des valeurs cibles sont proposées tableau R.

Tableau R : Proposition et recommandations pour la fixation des valeurs cibles

	Indicateurs spécifiques	Valeurs cibles proposées
IS 1	Relatifs au respect des spécifications générales d'enregistrement	100% *
IS 2	Délais	A fixer selon les objectifs du dispositif, et la réalité de terrain (le temps des analyses notamment, qui n'est pas réductible)
	Tous les autres indicateurs relatifs à l'IS 2	Moyennes nationales, 98%, 100%...

* Les critères de respect des spécifications générales d'enregistrement sont censés attester que l'EES existe effectivement, et constituent donc un préalable indispensable à toute analyse des résultats de la surveillance.

→ Compléter la section 6 – Indicateurs synthétiques

Exemple d'évaluation : le dispositif de déclaration des avortements		
<p>Les valeurs cibles ont été définies à partir des consignes existantes (formelles ou non), des objectifs du dispositif et des contraintes de terrain. Pour les validités de la durée de gestation et de l'identifiant bovin ainsi que pour la cohérence entre la date d'avortement et la date de visite, 2% d'erreurs a été considéré comme acceptable du fait du facteur humain entrant en jeu.</p> <p>Au total, 26 indicateurs ont donc été retenus pour attribuer les IS départementaux (Tableau S).</p> <p>Pour chaque département, le dénominateur de l'IS correspondait au nombre d'indicateurs calculés, qui pouvait être inférieur à 22 pour l'IS 2 : par exemple, si le format de la durée de gestation n'était jamais respecté aucune validité n'a été calculée. Pour faciliter la comparaison, dans les cas où les 22 indicateurs n'ont pu être calculés, l'IS 2 du département en question a été standardisé (par exemple un IS de 14/19 a été transformé en un IS de 16/22).</p> <p>Tableau S : Indicateurs et seuils d'acceptabilité retenus pour l'attribution des indicateurs synthétiques départementaux</p>		
	Indicateur	Valeur cible
IS 1 (/4)	Un bovin par intervention	100%
	Une intervention par avortement	100%
	Un RAI complet par intervention	100%
	Cohérence entre la date de visite et la date de RAI	100%
IS 2 (/22)	Délai moyen entre les dates d'avortement et de validation	15 jours
	Délai moyen entre les dates de validation et d'envoi du RAI	7 jours
	Complétude de l'identifiant bovin et de l'interprétation des résultats	100%
	Complétude de la date d'avortement et de la durée de gestation	Moyennes nationales
	Format de l'identifiant bovin, de la date d'avortement et de la durée de gestation	100%
	Validité de la durée de gestation et de l'identifiant bovin	98%
	Validité de la nature du prélèvement, de la matrice analysée, de la méthode et du résultat	100%
	Cohérence entre la date d'avortement et la date de visite	98%
Les six autres cohérences	100%	
<p>Les cohérences sont celles entre la nature du prélèvement la matrice analysée, entre la matrice analysée, l'analyte recherché et la méthode d'analyse, entre le MAM et le motif de non analyse, entre le MAM et la donnée analyse de confirmation, entre les</p>		

différentes analyses d'un RAI, entre les résultats et l'interprétation des résultats, et entre la date de l'évènement et la date de visite.

PRESENTATION DES RESULTATS

1.7 Résultats de l'évaluation qualitative

Les notes peuvent être présentées sous la forme d'un tableau récapitulatif, en mettant en évidence différents niveaux selon les notations (par exemple distinguer les notes supérieures à 2,5/3, celles entre 1,5/3 et 2,5/3 et celles inférieures à 1,5/3.

Exemple d'évaluation : le dispositif de déclaration des avortements

En ce qui concerne la DA, les résultats de l'évaluation qualitative ne concernent que les trois départements visités. Les moyennes des notes attribuées pour évaluer la pertinence et la crédibilité de chacune des données collectées sont résumées dans le Tableau T.

Tableau T : Moyennes des notes attribuées aux dimensions de second niveau par les acteurs du dispositif des trois départements visités (Rhône, Puy-de-Dôme, Haute-Savoie)

	Pertinence					Crédibilité				
	VS	L	D	G	Tot	VS	L	D	G	Tot
Date d'intervention	2,9	2,5	3,0	2,5	2,8	2,8	2,3	2,3	2,5	2,4
Numéro d'ordre	2,8	2,0	3,0	3,0	2,7	3,0	2,5	3,0	2,0	2,7
Identifiant de l'exploitation – atelier	3,0	3,0	3,0	3,0	3,0	3,0	2,3	3,0	2,5	2,6
Identifiant bovin	3,0	3,0	2,7	3,0	2,9	2,9	2,5	2,3	2,0	2,5
Date d'avortement	2,9	2,5	3,0	3,0	2,8	2,4	1,8	2,0	2,0	2,0
Durée de gestation	2,8	0,5	1,0	3,0	1,9	2,1	2,3	1,5	2,0	1,9
Matrice prélevée	2,8	3,0	3,0	3,0	2,9	3,0	2,8	3,0	2,0	2,7
Motif de non analyse	2,8	3,0	2,0	NA	2,7	2,7	3,0	3,0	NA	2,6
Etat du plan d'analyse	2,0	2,7	1,7	NA	2,1	3,0	3,0	3,0	NA	2,6
Matrice analysée	3,0	3,0	2,0	NA	2,8	3,0	3,0	3,0	NA	2,7
Analyte	3,0	3,0	1,5	NA	2,7	3,0	3,0	3,0	NA	2,7
Méthode	3,0	3,0	1,5	NA	2,8	3,0	3,0	3,0	NA	2,7
Analyse de confirmation	2,8	3,0	1,5	NA	2,6	3,0	3,0	3,0	NA	2,7
Résultat	3,0	3,0	3,0	3,0	3,0	3,0	3,0	3,0	3,0	2,8
Date de validation des résultats	1,7	2,8	1,0	3,0	2,0	3,0	3,0	2,0	3,0	2,7
Date du RAI	1,8	1,7	2,0	3,0	1,9	3,0	3,0	2,0	3,0	2,4
Etat du RAI	1,3	3,0	1,5	3,0	2,0	3,0	3,0	3,0	3,0	2,5
Laboratoire et laboratoire sous-traitant	2,0	3,0	2,5	3,0	2,5	3,0	3,0	2,5	3,0	2,7
Interprétation	2,8	2,0	1,5	NA	2,4	3,0	2,5	3,0	NA	2,6

VS = vétérinaire sanitaire, L = laboratoire, D = DDecPP, G = GDS, Tot = moyenne globale

Cases en blanc : note $>2,5/3$; cases en gris clair : $1,5/3 < \text{note} \leq 2,5/3$; cases en gris foncé : note $\leq 1,5/3$

La pertinence et la crédibilité ont été jugées mauvaises par tous les acteurs pour la durée de gestation et la date d'avortement. Les données collectées par le VS ont été jugées peu crédibles par les GDS et les laboratoires. Les données relatives aux analyses n'ont pas été notées par les GDS ; elles ont été jugées très crédibles par tous les autres, mais peu pertinentes par les DDecPP. Les données relatives au RAI et à l'intervention ont été globalement jugées peu pertinentes et peu crédibles.

Par ailleurs, la compréhension des données a été jugée très bonne par l'ensemble des acteurs. Les entretiens avec les différents acteurs rencontrés lors des déplacements n'ont pas permis d'obtenir des résultats exploitables pour la dimension accès / utilisation des données car les données sont très peu utilisées et le nombre de réponses est donc faible. De ce fait, aucun résultat n'est présenté ici.

1.8 Résultats de l'évaluation quantitative

Pour les indicateurs spécifiques, les résultats doivent être présentés par indicateurs. Il est nécessaire de présenter d'une part les résultats nationaux (une phrase peut suffire) et d'autre part les résultats départementaux sous la forme de box plot, dont le principe est présenté dans la figure E.

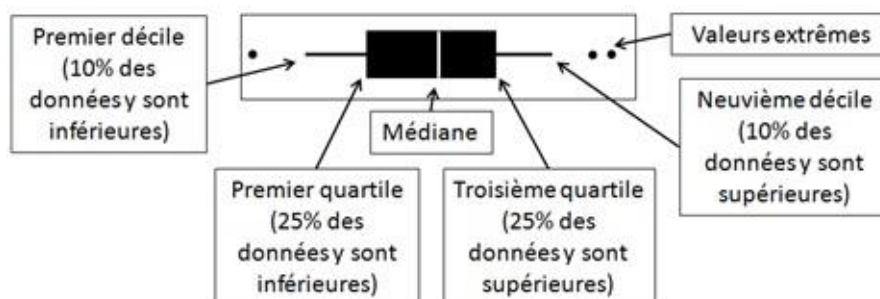


Figure E : Modalités d'interprétation d'un box plot

Les indicateurs synthétiques peuvent être présentés sous la forme d'un diagramme en barre ordonné par ordre croissant. Des cartographies peuvent également être faites afin de visualiser les indicateurs synthétiques, ainsi que le nombre d'entités élémentaires de surveillance par département et les indicateurs spécifiques non pris en compte dans le calcul des indicateurs synthétiques.

Exemple d'évaluation : le dispositif de déclaration des avortements

A – Indicateurs associés aux critères de sélection des déclarations d'avortements respectant les spécifications générales d'enregistrement

Des résultats ont été obtenus pour chaque département et pour l'ensemble du territoire. Ils sont présentés figure F, sous forme de box plot.

Les résultats de l'évaluation des indicateurs utilisés pour la sélection de DA respectant les spécifications générales d'enregistrement ont montré globalement une variabilité assez faible entre départements, même si ces indicateurs étaient faibles à très faibles pour certains (jusqu'à cinq départements entre 0 et 90%). Les indicateurs calculés au niveau national pour les critères A à D étaient respectivement de 99,9%, 96,5%, 98,7% et 99,6%.

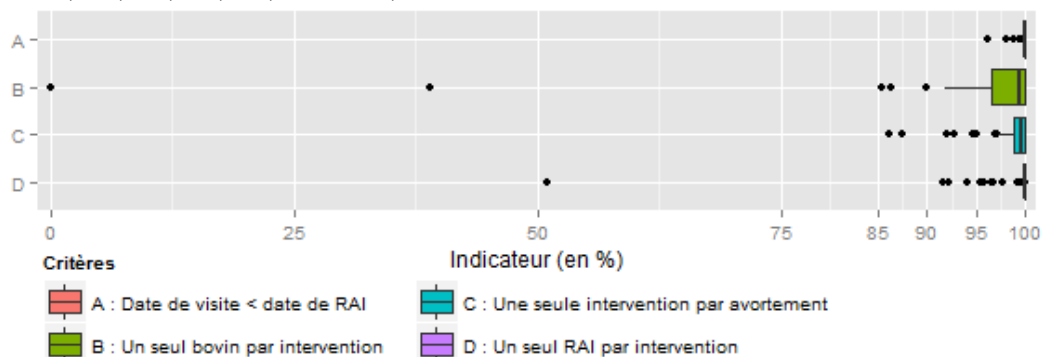


Figure F : Distribution des indicateurs témoignant de la proportion de DA considérées respectant les spécifications générales d'enregistrement et calculés à l'échelle départementale

B- Indicateurs associés aux dimensions de premier niveau pour les déclarations d'avortements respectant les spécifications générales d'enregistrement

Comme pour les critères de sélection des DA respectant les spécifications générales d'enregistrement, des résultats ont été obtenus pour chaque département et pour l'ensemble du territoire. Ils sont présentés figures G à K sous forme de boxplot.

1 – Complétude, format et validité de certaines données (analyse par « colonne »)

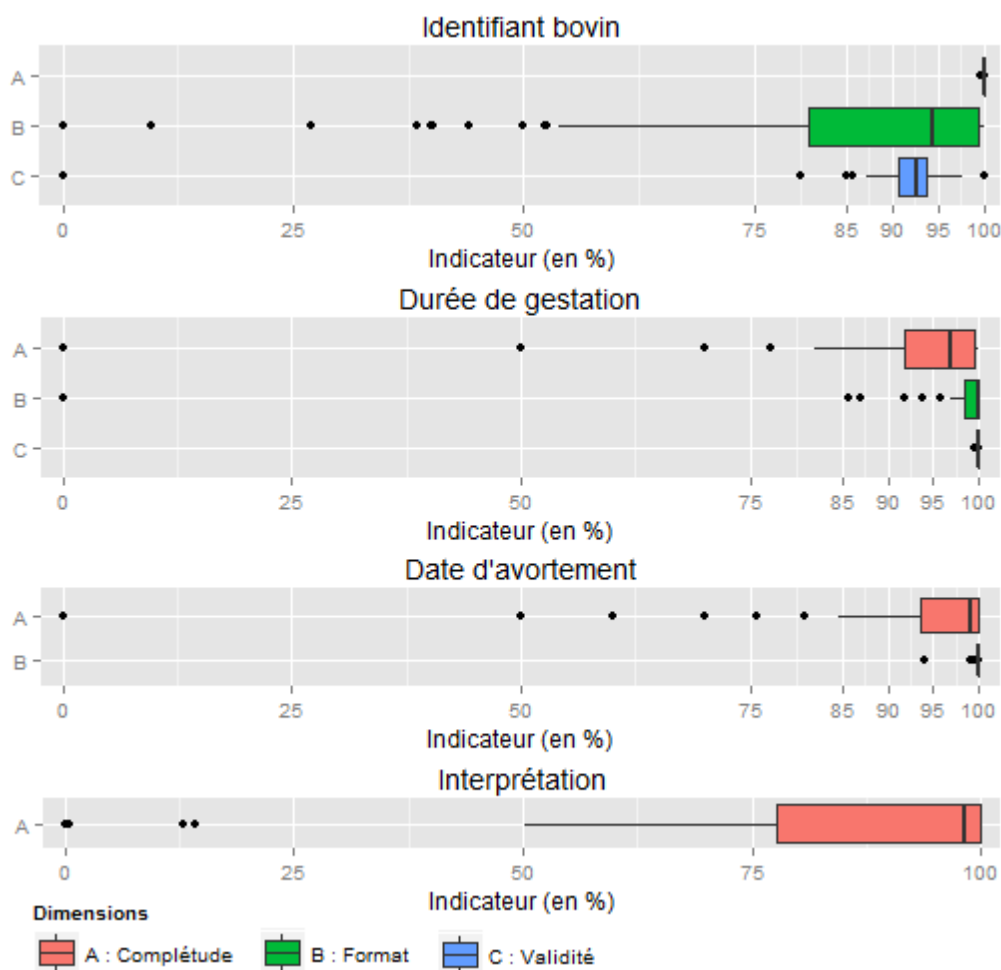


Figure G : Distribution des indicateurs témoignant de la qualité des données pour l'identifiant bovin, la durée de gestation, la date d'avortement et l'interprétation, calculés à l'échelle départementale

L'identifiant bovin était quasiment toujours renseigné (complétude élevée, avec un indicateur national de 99,98%). Le format était plus ou moins bien respecté en fonction des départements (indicateur national de 87,7%, et indicateur départemental variant entre 80,9 et 99,5% pour trois quarts des départements, en-dessous de 50% pour sept départements). L'identifiant bovin, lorsqu'il était au bon format, était valide pour 92,6% des bovins, avec une faible variabilité interdépartementale. Ainsi, au niveau national et parmi les DA respectant les spécifications générales d'enregistrement, 81,0% des analyses enregistrées dans SIGAL avaient un identifiant bovin valide ($0,926 \times 0,877 \times 0,998 = 0,810$).

Pour la durée de gestation, la complétude était assez élevée mais assez variable (indicateur national de 95,2%, et indicateur départemental situé entre 91,9 et 99,4% pour trois quarts des départements). Le respect du format était toujours élevé (indicateur national de 99,0%) et peu variable, avec tout de même quelques départements pour lesquels il était moins bien respecté (inférieur à 95%). Lorsque la durée de gestation était au bon format, elle était valide pour 99,95% des bovins au niveau national, avec une très faible variabilité interdépartementale (99,4% à 100%).

Pour la date d'avortement, la complétude était globalement similaire à celle de la durée de gestation (indicateur national de 97,3%, indicateur départemental situé entre 93,6 et 99,9% pour 3/4 des départements) ; le format était quasiment toujours respecté (indicateur national de 99,9%).

Pour l'interprétation des résultats par le laboratoire, la complétude avait une médiane assez élevée (98,1%) mais une très forte variabilité (indicateur départemental variant entre 77,6 et 99,96% pour trois quarts des départements). L'indicateur national était de 88,8%.

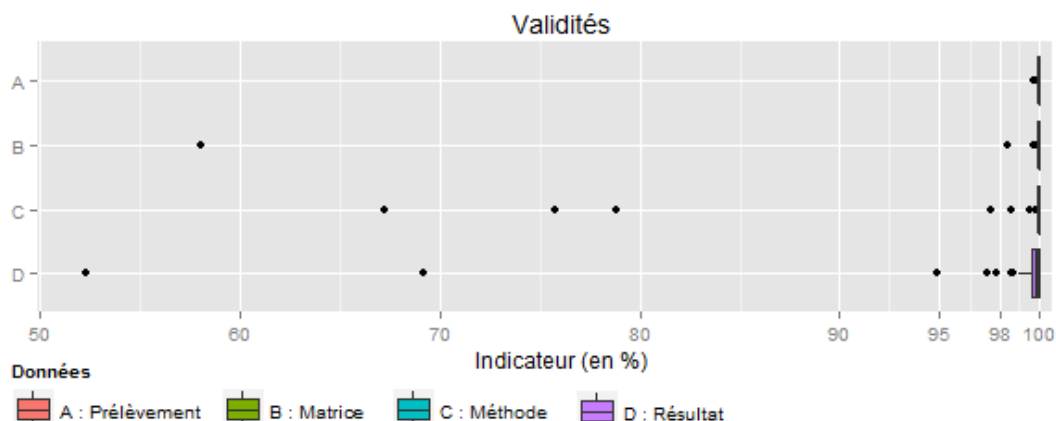


Figure H : Distribution des validités du prélèvement, de la matrice, de la méthode et du résultat, calculées à l'échelle départementale

La validité de la nature du prélèvement était excellente : l'indicateur départemental était de 100% pour 80 départements. Les validités de la matrice, de la méthode d'analyse et des résultats étaient également très bonnes. Les indicateurs départementaux associés à la validité de la matrice et de la méthode d'analyse étaient de 100% dans respectivement 80 et 74 départements ; la médiane de l'indicateur départemental associé à la validité du résultat était de 99,9%. Malgré cela, quelques départements se démarquaient fortement avec des indicateurs variant entre 50 et 80%. Les indicateurs nationaux des validités de la nature du prélèvement, de la matrice, de la méthode et du résultat étaient respectivement de 99,98%, 99,6%, 96,9% et 97,9%.

2 – Vitesse de circulation des données

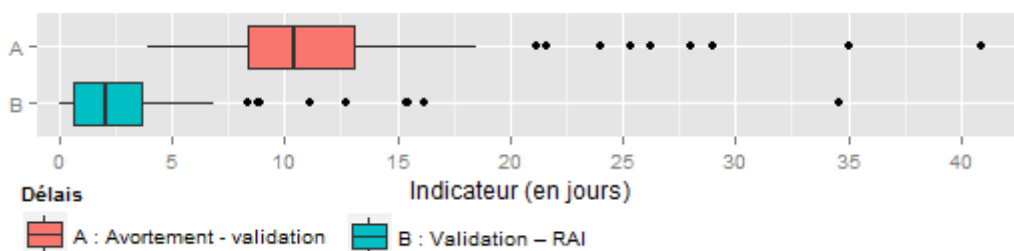


Figure I : Distribution des indicateurs témoignant de la vitesse de circulation des données calculés à l'échelle départementale (données tronquées à 43 jours)

L'évaluation de la vitesse de circulation des données a montré que les deux indicateurs avaient une variabilité importante entre les départements. Par ailleurs, le délai moyen entre l'avortement et la validation des résultats de la Charente n'apparaît pas sur la figure I : il était de presque 122 jours. Les indicateurs nationaux des délais A et B étaient respectivement de 11,7 jours et de 2,9 jours.

3 – Validité et complétude des déclarations d'avortements (analyse « par ligne »)

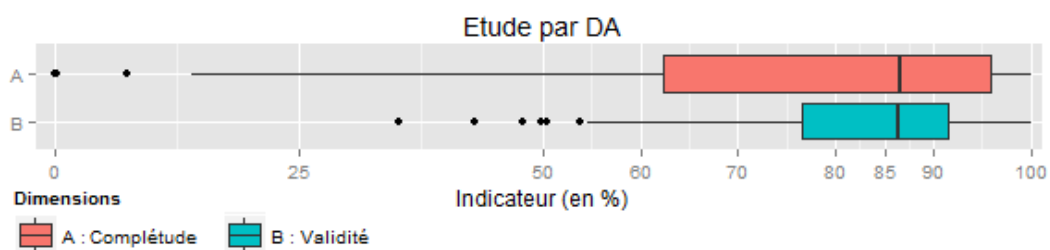


Figure J : Distribution de la complétude et de la validité des déclarations d'avortements calculées à l'échelle départementale

L'évaluation de la qualité des données par DA a montré une très forte variabilité de la complétude (pour trois quarts des départements, de 62,4 à 95,8% des DA étaient complètement renseignées) et de la validité (pour trois quarts des départements, de 76,6 à 91,6% des DA complètement renseignées l'étaient avec des données valides). Par ailleurs, les médianes de ces deux indicateurs départementaux étaient relativement basses, proches de 85%. Les indicateurs nationaux de la complétude et de la validité étaient respectivement de 83,6% et 80,7%. Ainsi seulement 67,5% des DA respectant les spécifications générales d'enregistrement étaient entièrement valides (83,6*80,7).

4 – Cohérence entre les données

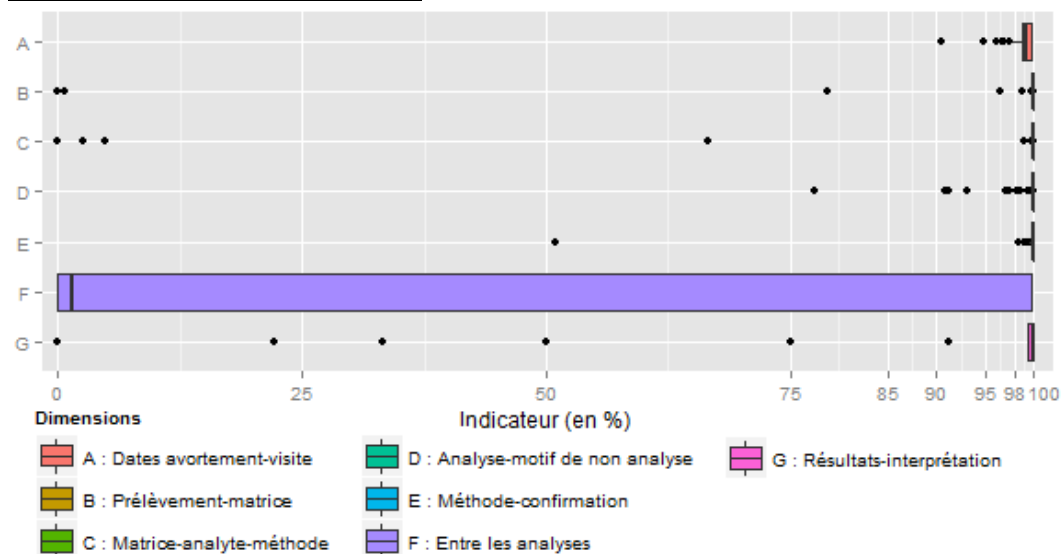


Figure K : Distribution des indicateurs témoignant de la cohérence entre les données, calculés à l'échelle départementale

La cohérence entre les dates d'avortement et de visite du vétérinaire était bonne et assez peu variable entre les départements. Les cohérences entre la nature du prélèvement et la matrice analysée, entre la matrice analysée, l'analyte et la méthode, entre l'analyse et le motif de non analyse, et entre la méthode et la donnée « confirmation » étaient excellentes et avec une variabilité très faible ; toutefois, pour quelques départements, ces indicateurs étaient inférieurs à 85%. Les indicateurs nationaux des cohérences A à E étaient respectivement de 99,0%, 96,3%, 97,6%, 98,6% et 99,7%.

La cohérence entre les différentes analyses réalisées par le laboratoire avait une variabilité très élevée : les différentes analyses réalisées étaient cohérentes pour moins de 5% des RAI dans 48 départements, et pour plus de 95% des RAI dans 32 départements. La médiane de l'indicateur départemental et l'indicateur national étaient très faibles (respectivement 1,5% et 46,5%).

La cohérence entre les résultats d'un RAI et leur interprétation était très bonne (médiane de 100% et indicateur national de 99,7%) et assez peu variable, mais quinze départements se démarquaient fortement (avec un indicateur inférieur à 92%, dont neuf départements avec un indicateur de 0%).

Indicateurs synthétiques départementaux

Les dispersions des IS départementaux (IS 1 et IS 2) sont présentées sur les Figures L et M.

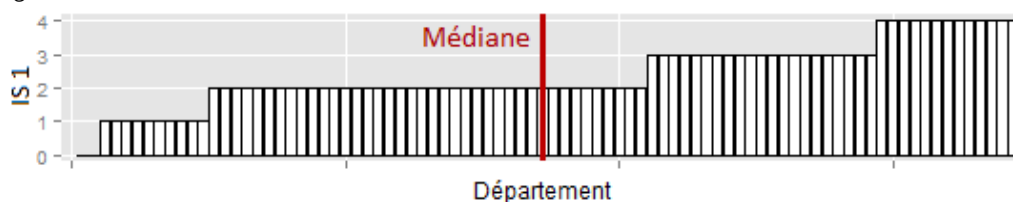


Figure L : Distribution de l'indicateur synthétique 1 attribué en fonction des valeurs cibles retenues

Les IS 1 sont centrés sur 2/4. L'Aisne et la Haute-Loire avaient les IS les plus bas (0/4), treize départements avaient un IS de 4/4.

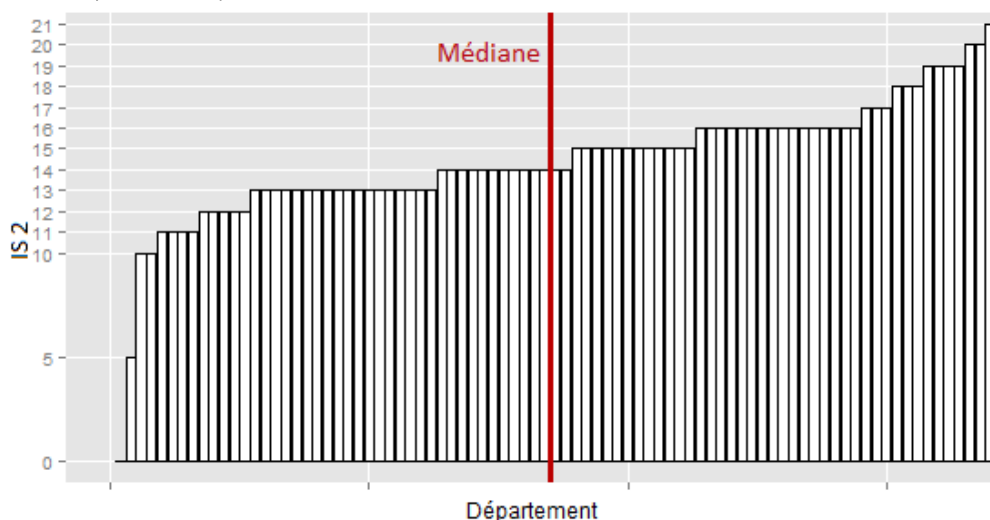
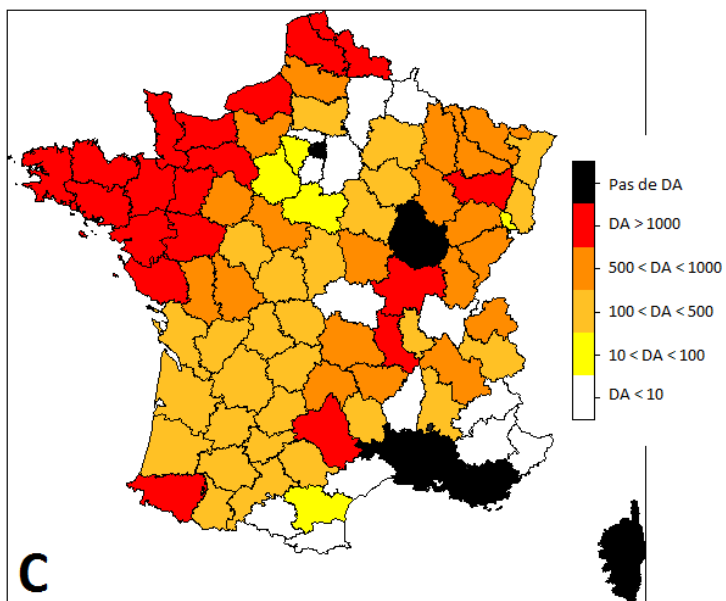
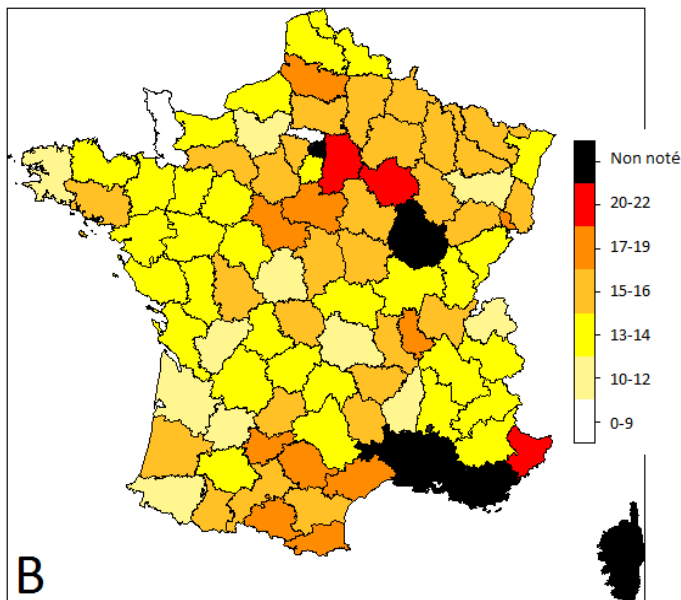
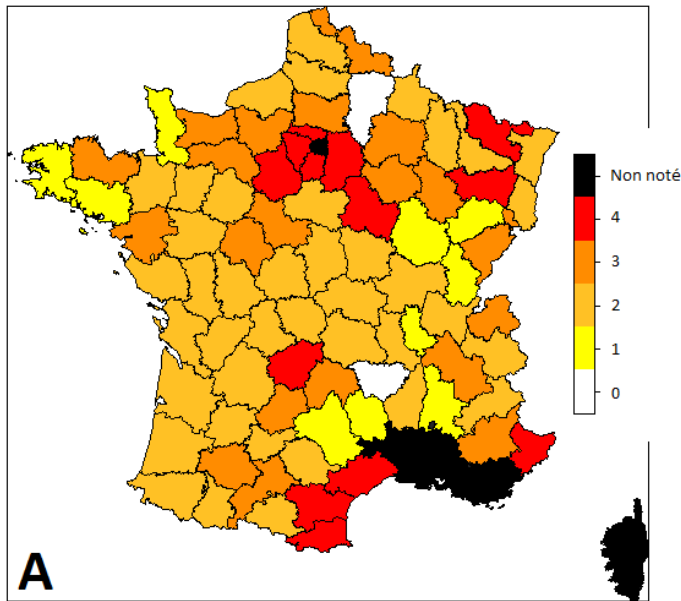


Figure M : Distribution de l'indicateur synthétique 2 attribué en fonction des valeurs cibles retenues

Très peu de départements avaient un IS 2 vraiment faible (seuls le Val d'Oise (0/22) et la Manche (5/22)) ou vraiment élevée (la Seine-et-Marne (21/22), l'Aube et les Alpes-Maritimes (20/22)). La grande majorité des départements étaient entre 13 et 16/22. Toutefois, 15% des départements (soit 13 départements sur 86) étaient au dessus de 16/22.

Ces IS n'ont pas été pondérés en fonction du nombre de DA de chaque département. Toutefois, la Figure N présente le nombre de DA par département, afin de pouvoir relativiser l'impact des IS départementaux sur la qualité nationale des données de DA. Les distributions géographiques des IS départementaux sont également présentées en Figure N celles de la complétude et de la validité par DA en Figure O.



Les départements en noir sont ceux pour lesquels aucun IS n'a pu être attribué, c'est-à-dire pour l'IS 1 : aucune DA n'était enregistrée dans SIGAL
pour l'IS 2 : aucune DA n'était enregistrée dans SIGAL ou aucune DA respectait les critères de sélection des DA ayant effectivement eu lieu (cas de la Côte d'Or).

Figure N : Distribution géographique A : de l'indicateur synthétique 1, B : de l'indicateur synthétique 2, C : du nombre de DA respectant les spécifications générales d'enregistrement.

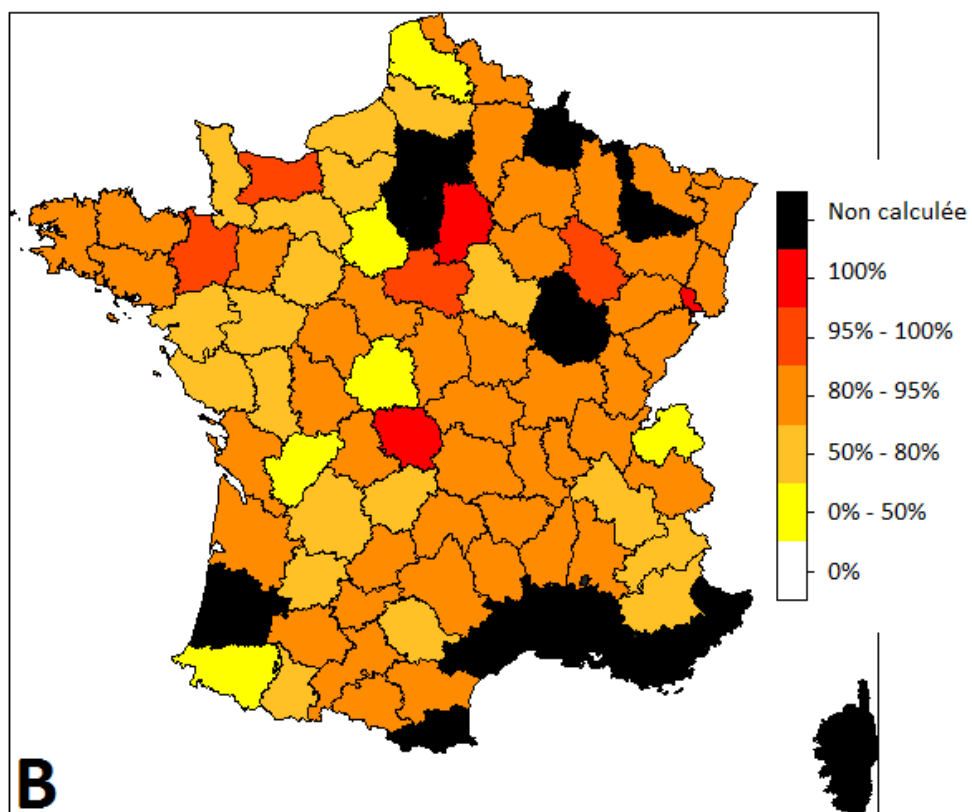
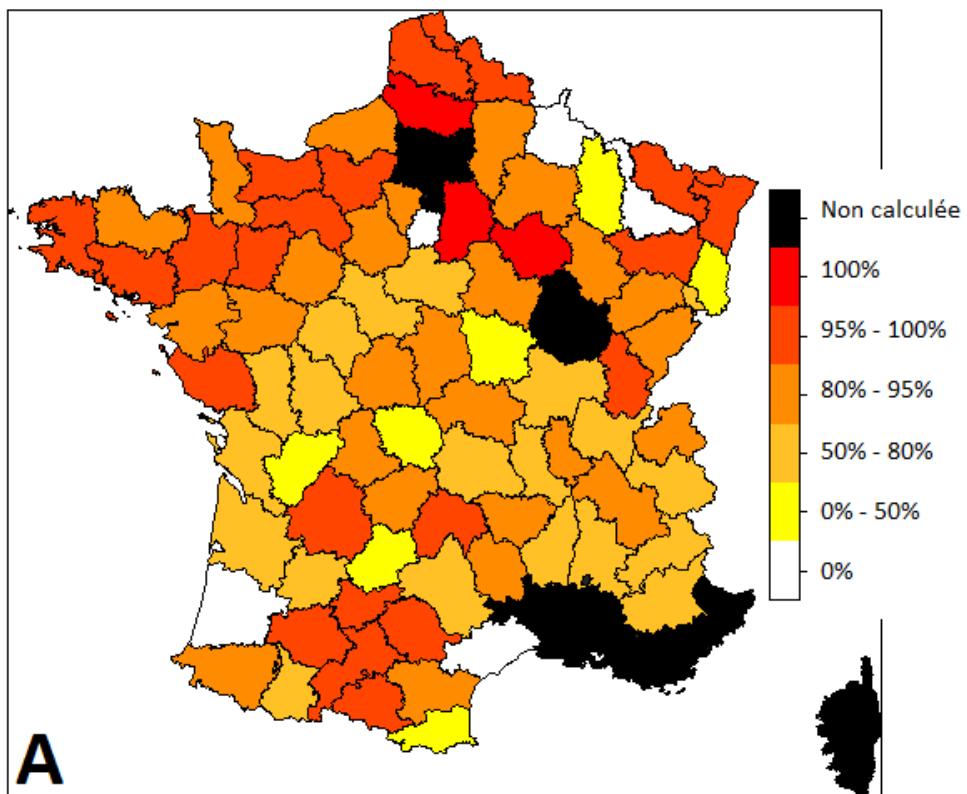


Figure O : Distribution géographique A : de la complétude et B : de la validité des déclarations d'avortements.

D'après les cartes des figures N et O, il ne semble pas y avoir de modèle géographique entre les IS et le nombre de DA, ni entre la validité et la complétude des DA.

RECHERCHE DE CORRELATIONS (SECTION 7 DU FORMULAIRE)

La recherche de corrélations significatives (identifiées avec le test de Pearson) a pour but de s'assurer que les résultats des indicateurs spécifiques sont indépendants les uns des autres, et qu'ils peuvent être interprétés séparément.

Les corrélations à étudier sont celles entre :

- les indicateurs d'une même donnée (par exemple entre la complétude et le format d'une donnée) ;
- les indicateurs d'une même dimension (par exemple entre les validités de deux données) ;
- les indicateurs synthétiques et le nombre d'EES par département.

→ Compléter la section 7 – Recherche de corrélations

Exemple d'évaluation : le dispositif de déclaration des avortements

Plusieurs indicateurs n'ont pas été inclus dans la recherche de corrélations :

- ceux témoignant de la complétude de l'identifiant bovin, de la validité de la durée de gestation et du format de la date d'avortement : ces indicateurs étaient très bons et très peu variables entre les départements ;
- ceux témoignant des validités du prélèvement, de la matrice, de la méthode et du résultat : ces indicateurs étaient très bons, peu variables et les seuls départements qui se démarquaient avec des valeurs extrêmes n'étaient pas les mêmes selon les indicateurs.

Les corrélations étudiées peuvent être regroupées en plusieurs catégories. Ce sont celles entre :

- indicateurs d'une même donnée. C'est par exemple la corrélation entre le format et la validité de l'identifiant bovin ;
- indicateurs d'une même dimension. C'est par exemple la corrélation entre deux des cohérences analysées ;
- les IS et le nombre de DA dans un département ;
- l'IS 1 et l'IS 2.

Sur les 38 corrélations testées, dix se sont avérées significatives : le test de Pearson était significatif ($p\text{-value} < 0,5$). Parmi elles, huit étaient positives et deux négatives. Les corrélations significatives sont présentées dans le Tableau U.

Tableau U : Types de corrélations pour lesquelles le test de Pearson était significatif

Corrélation étudiée		p-value	Coefficient de corrélation		
Indicateurs d'une même donnée	Identifiant bovin	Format / validité	$3,6 \times 10^{-4}$	0,39	
	Durée de gestation	Complétude / format	$4,1 \times 10^{-11}$	0,65	
Indicateurs d'une même dimension	MCD	Cohérence entre la date de la visite et la date du RAI / unicité du bovin par intervention	$6,5 \times 10^{-11}$	0,63	
	Complétude	Date d'avortement / durée de gestation	$3,4 \times 10^{-09}$	0,59	
	Format	Identifiant bovin / durée de gestation	$9,0 \times 10^{-07}$	0,52	
	Cohérence		Dates d'avortement – date de visite / matrice – analyte – méthode	$2,9 \times 10^{-03}$	0,32
			Prélèvement – matrice / méthode – confirmation	$5,2 \times 10^{-08}$	0,56
			Matrice – analyte – méthode / les analyses d'un RAI	$4,3 \times 10^{-02}$	-0,22
		Les analyses d'un RAI / résultats – interprétation	$9,9 \times 10^{-08}$	0,62	
IS et nombre de DA		IS 2 / nombre DA	$1,4 \times 10^{-02}$	-0,27	

DISCUSSION ET INTERPRETATION DES RESULTATS

Les résultats obtenus doivent être discutés et interprétés à la lumière des résultats de l'évaluation qualitative, des entretiens avec les différents acteurs et en tenant compte des éventuelles corrélations mises en évidence.

Cela doit permettre de mettre en évidence des facteurs influençant la qualité des données et à partir de là de proposer des axes d'améliorations.

Par ailleurs, la discussion doit permettre de s'assurer que les objectifs de l'évaluation fixés avant sa mise en œuvre ont bien été atteints.

1.9 Facteurs influençant la qualité des données

Les facteurs influençant la qualité des données peuvent être ordonnés dans trois catégories afin de mieux identifier les axes d'amélioration qui en découlent :

- matériel : il s'agit des différents outils utilisés pour la gestion des données dans son ensemble (collecte, saisie, validation) et dans le fonctionnement du dispositif ;
- méthode : il s'agit des différentes procédures, formelles ou non qui peuvent exister ;
- acteurs : il s'agit des facteurs qui dépendent de l'influence des opérateurs dans la gestion des données mais aussi de l'organisation des acteurs pour mettre en œuvre les modalités de surveillance correspondant au dispositif évalué.

Exemple d'évaluation : le dispositif de déclaration des avortements

L'évaluation de la qualité des données du dispositif de DA avait pour but d'évaluer de façon approfondie cette qualité. Elle devait permettre de mettre en

évidence les éventuels défauts de qualité des données, leurs causes principales et de proposer des pistes pour l'améliorer.

Pour cela, différents indicateurs ont été construits et calculés, les départements ont été notés à partir des résultats de ces indicateurs et des entretiens ont été conduits avec les acteurs du dispositif pour prendre en compte leur perception de la qualité des données.

Malgré certaines limites de la méthode utilisée (non détaillées dans ce guide) cette évaluation permet de disposer d'un premier état des lieux de la qualité des données du dispositif de DA. Les résultats de l'évaluation de la qualité des données de DA sont globalement satisfaisants et il y a peu de départements pour lesquels la qualité des données était vraiment faible. Toutefois, une marge de progrès existe :

- à l'échelle nationale, deux tiers des DA (soit 41 420 DA sur 63 073) étaient entièrement complètes et valides
- les IS 2 départementaux étaient regroupés autour de 14/22, avec une marge de progression clairement mise en évidence : 15% de départements avaient un IS 2 supérieur à 16/22.

Trois points particulièrement préoccupants ont été mis en évidence : ils peuvent impacter la qualité de l'évaluation du dispositif et la connaissance de la population surveillée. L'identifiant bovin était non valide dans près d'une DA sur cinq (en ne comptant que les DA respectant les spécifications générales d'enregistrement), alors qu'il permet de s'assurer que l'avortement a bien eu lieu. Le délai d'obtention des résultats était *a priori* trop élevé dans un nombre important de départements, ce qui limite les éventuelles corrections qui seraient à apporter dans les données mal ou non renseignées. Les types d'analyses réalisées et leur ordre de réalisation préconisé dans la fiche de plan n'étaient pas respectés pour 80% des DA, ce qui traduit un dysfonctionnement dans la mise en œuvre du dispositif.

Enfin, une disparité interdépartementale très importante a été observée, du fait de l'absence d'harmonisation nationale des modalités de collecte, de saisie et de validation des données, ainsi que de la diversité des outils informatiques et de leur paramétrage, notamment au niveau des Lims.

Il s'agit à présent d'interpréter les résultats obtenus, de présenter les facteurs influençant la qualité des données et de proposer des perspectives d'amélioration.

Critères de sélection des déclarations d'avortements respectant les spécifications générales d'enregistrement

Les résultats de l'étude de la cohérence entre la date de visite et la date de RAI étaient plutôt satisfaisants, et les écarts vis-à-vis de ce qui était attendu peuvent être expliqués principalement par des erreurs de saisie, des dates peu lisibles sur la feuille de commémoratifs et l'absence de validation de la donnée.

La faible valeur obtenue pour l'indicateur « deux interventions pour un avortement » dans certains départements s'explique en partie par des règles de gestion locales et par des disparités dans les droits de suppression : comme vu dans la description du dispositif, ces droits sont attribués localement. Les interventions sont généralement créées en masse à partir des RAI exportés dans SIGAL bien que certains RAI soient erronés ou incomplets. Ensuite, les interventions créées à partir de ces RAI ne sont pas toujours supprimées, notamment lorsque les droits de suppression n'ont pas été attribués à la personne gérant les interventions.

L'indicateur « une intervention ne concerne qu'un bovin » était le moins satisfaisant. On peut distinguer deux types de justifications :

- soit il s'agit d'un choix délibéré de la DDecPP :

o afin de ne payer qu'une visite lorsque par exemple deux DA sont faites par le même VS au même endroit le même jour, une seule intervention est créée : il s'agit alors de doublons ponctuels, ou

o pour distinguer (en Côte d'or notamment) les différents prélèvements sur un même animal : cet écart concernait la grande majorité des DA car les interventions comprenaient presque toujours deux prélèvements, identifiés de manière distincte (par exemple l'identifiant du bovin +/- abréviation EC pour écouvillon). Ces interventions ne concernaient en réalité qu'un bovin, mais la saisie de l'identifiant était non conforme ;

- soit il s'agit du seul choix du laboratoire : un VS effectue deux DA le même jour et au même endroit (en renseignant une ou deux feuilles de commémoratifs), et le laboratoire choisit de n'exporter qu'un seul RAI.

Les interventions auxquelles plusieurs RAI sont rattachés s'expliquent également par des règles de gestion locales, pour ne payer qu'une visite lors de plusieurs DA le même jour, au même endroit, par le même VS.

Vitesse de circulation des données

Le délai entre la date d'avortement et la date de validation des résultats était supérieur à 20 jours dans 10% des départements étudiés. Ce délai tient compte de différentes étapes, qui impliquent successivement l'éleveur, le VS et le laboratoire : l'appel de l'éleveur, la réalisation de la visite (influence entre autre de l'étendue de la clientèle et de l'organisation des VS), l'envoi des prélèvements (influence de l'existence ou non d'une navette), la réalisation des analyses, l'obtention des résultats et la validation des résultats (influence de l'organisation interne du laboratoire).

Par ailleurs, ce délai entre la date d'avortement et la date de validation des résultats peut être élevé sans pour autant remettre en cause la réactivité du dispositif : l'organisation de certains laboratoires est telle que les résultats ne sont validés que lorsque toutes les analyses sont réalisées, notamment celles relatives au diagnostic différentiel des avortements (qui peuvent être plus longues ou réalisées moins fréquemment que les analyses relatives à la brucellose), et sous réserve que l'une des personnes habilitée pour valider les résultats soit disponible. Ainsi, une évaluation du délai réel d'obtention des résultats devrait reposer sur l'analyse du délai entre la date d'avortement et la date d'obtention des résultats liés à la brucellose. Dans ce cas, il serait possible d'avoir une estimation de la réactivité réelle du dispositif, en faisant l'hypothèse que la DDecPP est informée dès l'obtention d'un résultat sérologique positif vis-à-vis de la brucellose par le laboratoire.

Enfin, d'après les résultats des indicateurs relatifs aux dimensions de second niveau, la fiabilité de la date d'avortement est assez moyenne. En effet, dans certains départements, si la date d'avortement n'était pas renseignée sur la feuille de commémoratifs, elle était saisie dans le système informatique par défaut comme identique à la date de visite. Ainsi, dans un certain nombre de cas, l'indicateur du délai d'obtention des résultats a été surestimé (le délai entre la date d'avortement et la date de visite n'est pas pris en compte).

Données relatives aux commémoratifs collectés par le VS

→ **Identifiant bovin**

Près d'une analyse sur cinq était renseignée avec un identifiant bovin absent, au mauvais format ou non valide. Pourtant, cette donnée permet de valider la survenue effective de l'avortement et d'assurer une traçabilité indispensable en cas de résultat positif. Les écarts observés peuvent être expliqués par :

- Un problème au niveau de la collecte : une donnée non collectée ou collectée seulement partiellement (par exemple uniquement le numéro de travail), une donnée illisible ;
- Un problème postérieur à la collecte : une erreur de saisie, une absence de contrôle local (au mieux sur la validité du numéro, au minimum sur le format).

Il est important de considérer les trois indicateurs ensemble (complétude, format et validité), car ils reflètent des dysfonctionnements différents. En effet, certains départements renseignent obligatoirement l'identifiant dans un bon format (du fait par exemple de contraintes informatiques locales) sans pour autant vérifier sa validité (si seul le numéro de travail est connu, le reste peut être remplacé par des 0 ou des X).

→ **Durée de gestation**

Les indicateurs des dimensions de premier niveau se rapportant à la durée de gestation étaient globalement satisfaisants. Les faibles écarts peuvent s'expliquer par le fait que les consignes ne sont pas toujours claires et/ou complètes (par exemple, certaines durées de gestation étaient renseignées en semaines).

De plus, cette donnée apparaît peu pertinente et peu fiable pour les acteurs rencontrés : ils ne sont donc pas enclin à passer du temps à sa collecte, sa saisie ou à procéder à une validation. Les résultats des indicateurs de premier niveau sont donc à nuancer : malgré une validité *a priori* élevée (estimée comme la proportion de DA pour lesquelles la durée de gestation est un nombre entre 0 et 10), il est probable que cette donnée ne reflète pas si bien la réalité.

→ **Date d'avortement**

La complétude de la date d'avortement était plutôt satisfaisante. Toutefois, comme cela a été dit précédemment, elle peut être saisie par défaut comme identique à la date de visite : la complétude est augmentée mais avec une fiabilité diminuée. Cette remarque va de pair avec les faibles notes attribuées à cette donnée par les différents acteurs (dimensions de second niveau), qui ont en effet peu confiance en l'exactitude de la date d'avortement.

Données relatives aux analyses réalisées par le laboratoire

La qualité de certaines de ces données est liée à l'existence et à la qualité de la programmation du Lims du laboratoire. Les dimensions concernées sont la complétude de l'interprétation, les validités de la matrice, de la méthode et du résultat et les cohérences entre le prélèvement et la matrice, entre la matrice, l'analyte et la méthode, entre l'analyse et le motif de non analyse, entre l'analyse et la donnée confirmation et entre les résultats et leur interprétation. Les indicateurs relatifs à ces dimensions sont très bons si le paramétrage est bien fait, très mauvais s'il est mal fait et moyens s'il n'est que partiel (par exemple l'interprétation n'est prévue que pour certains résultats), ce qui explique les différentes allures de la

majorité des box plots (résultats très homogènes ou au contraire très étendus). Les résultats de ces indicateurs attestent d'un paramétrage globalement satisfaisant, excepté pour la complétude de l'interprétation.

Par ailleurs, la validité du résultat était un peu plus faible par rapport aux autres données pour lesquelles la validité a été étudiée. En effet, un résultat n'a été considéré valide que s'il était renseigné comme « positif » ou « négatif », conformément à la fiche de plan EBRUBVA. Ainsi, les résultats obtenus par la méthode de fixation du complément faisant apparaître le titre n'ont pas été considérés comme valides alors qu'ils le sont pour les laboratoires.

Pour la cohérence entre les différentes analyses d'un même RAI, la médiane de l'indicateur départemental et l'indicateur national étaient très différents : il y a 46,5% des RAI enregistrés au niveau national pour lesquels les analyses étaient cohérentes, mais dans la moitié des départements, il y avait moins de 1,5% des RAI pour lesquels les analyses étaient cohérentes. Cette différence s'explique en partie par le fait que la médiane correspond à la dispersion des départements (la moitié sont au-dessus et l'autre moitié en-dessous) alors que l'indicateur national prend en compte les données au niveau national : les départements avec beaucoup de DA ont plus de poids que ceux avec peu de DA. Elle s'explique également par le rôle du nombre de DA par département : si les départements avec beaucoup de DA ont un indicateur plus faible que celui des départements avec peu de DA, la différence entre la médiane et l'indicateur national est logiquement importante. Cependant, aucune corrélation n'a été mise en évidence entre cet indicateur et le nombre de DA.

Cet indicateur est pourtant un point critique car la réalisation d'analyses supplémentaires est à l'origine d'une dépense inutile pour l'Etat. Cet écart important s'explique par un non-respect des consignes de la fiche de plan et des recommandations du LNR. Ce non-respect fait probablement suite à une absence de mise à jour des pratiques : avant 2010, il était demandé de réaliser en première intention deux analyses sérologiques (une EAT et une FC ou une EAT et une ELISA), alors qu'une seule analyse est désormais demandée (Anonyme, 2008b).

Corrélations entre indicateurs

Dix corrélations significatives ont été mises en évidence : il est probable que certains facteurs influençant la qualité des données soient communs à plusieurs indicateurs. Ces résultats doivent être en partie considérés avec précaution car compte tenu de la multiplicité des tests, l'obtention d'un résultat significatif « par erreur », alors qu'il n'existe pas de réelle corrélation, est augmentée. Toutefois, des explications à la mise en évidence de certaines de ces corrélations peuvent être avancées.

Les corrélations relatives aux commémoratifs (format / validité de l'identifiant bovin, complétude / format de la durée de gestation, complétudes de la date d'avortement / de la durée de gestation, formats de l'identifiant bovin / de la durée de gestation) peuvent s'expliquer par une logique dans les gestions locales des commémoratifs : temps consacré, règles locales et implication des acteurs. Toutefois, les contraintes informatiques peuvent être indépendantes de l'existence d'une validation des données : par exemple, le format de l'identifiant bovin peut être

contraint informatiquement alors que la validité reflète plutôt l'existence effective d'une validation. Ceci pourrait expliquer que la corrélation ne soit pas plus marquée.

Les corrélations relatives aux données se rapportant aux analyses peuvent s'expliquer par le niveau de compréhension et de suivi de la fiche de plan dans son ensemble. Ces corrélations sont les suivantes :

- corrélation entre l'indicateur témoignant de la cohérence entre la nature du prélèvement et la matrice et l'indicateur témoignant de la cohérence entre la méthode et la donnée confirmation ;
- corrélation entre l'indicateur témoignant de la cohérence entre les analyses d'un RAI et l'indicateur témoignant de la cohérence entre les résultats et leur interprétation.

La corrélation négative entre l'IS 2 et le nombre de DA était faible. Une hypothèse principale est à envisager pour expliquer cette corrélation. Des départements avec peu de DA reçoivent moins de RAI et donc la création d'intervention « en masse » à partir de nombreux RAI est alors moins fréquente. De plus, il y a probablement moins d'automatismes et donc le temps passé sur chaque DA est plus long (chaque action doit être réfléchie). Ainsi, cela incite plus facilement à la validation des données, alors que le temps global passé sur la totalité des DA reste potentiellement plus faible.

La corrélation négative entre la cohérence matrice – analyte – méthode et la cohérence entre les différentes analyses d'un RAI est faible ainsi que la corrélation entre cohérence entre la date de visite et la date d'envoi du RAI / l'unicité du lien intervention – bovin n'ont pas pu être expliquées.

Par ailleurs, dans un certain nombre de cas, il n'y avait pas de corrélation, ou les corrélations mises en évidence étaient parfois faibles (avec un coefficient de corrélation inférieur à 0,4 pour quatre des dix corrélations). En effet, la qualité de chaque indicateur repose majoritairement sur des facteurs qui lui sont propres :

- elle dépend d'intervenants spécifiques (cas du délai entre l'avortement et la validation des résultats qui dépend des éleveurs, des VS et du laboratoire, et du délai entre la validation des résultats et l'envoi du RAI qui ne dépend que du laboratoire) ;
- elle dépend de contraintes informatiques qui peuvent être attribuées spécifiquement à chaque donnée de manière indépendante : ceci explique l'absence de corrélation entre la majorité des indicateurs dépendant de la programmation du Lims. Les laboratoires peuvent réaliser une programmation partielle, variable d'un département à l'autre. De plus, une programmation nécessite une mise à jour, qui n'est pas obligatoirement faite à la même fréquence et de la même manière dans les différents départements.

Ainsi, même si quelques corrélations ont été mises en évidence, on peut considérer que globalement chaque indicateur permet d'étudier différents aspects du dispositif.

Les facteurs influençant la qualité des données peuvent finalement être regroupés en trois catégories : le matériel, la méthode et les acteurs (Tableau V), à partir desquels différents axes d'amélioration peuvent être envisagés.

Tableau V : Facteurs influençant la qualité des données du dispositif de DA

Indicateur		Matériel	Méthode	Acteurs
DA respectant les spécifications générales	Date visite postérieure à date de RAI	Feuille de commémoratifs – pas de contrôle SIGAL	Absence ou erreur de consigne	Saisie – lisibilité
	Deux bovins par intervention			Temps consacré – collecte
	Deux RAI par intervention	Pas de contrôle SIGAL		
	Deux interventions par avortement	Droits de suppression		Erreur dans le RAI
Circulation des données	Délai entre date de visite et date de validation des résultats	Existence d'une navette, analyses groupées	Absence ou erreur de consigne	Motif d'appel, clientèle, temps consacré, intérêt, organisation
	Délai entre date de validation des résultats et date d'envoi du RAI	Paramétrage de l'exportation		Temps consacré, intérêt
Complétude	Identifiant bovin	Contrôle SIGAL défectueux	Absence ou erreur de consigne	Absence de collecte – lisibilité
	Durée de gestation	Donnée non connue		
	Date d'avortement			Contrôle
	Interprétation des résultats	Paramétrage du Lims		
Format	Identifiant bovin	Pas de contrôle SIGAL	Absence ou erreur de consigne	Collecte – lisibilité – contrôle
	Durée de gestation	Pas de contrôle SIGAL		Absence de contrôle
	Date d'avortement			
Validité	Identifiant bovin		Absence / erreur de consigne	Collecte – lisibilité – contrôle
	Durée de gestation	Pas de contrôle SIGAL		Contrôle
	Prélèvement			Saisie
	Matrice			
	Méthode	Paramétrage du Lims		
	Résultat			
Cohérences	Date de l'avortement et date de visite	Pas de contrôle SIGAL		Saisie – lisibilité
	Nature du prélèvement et matrice	Paramétrage du Lims		Contrôle
	Matrice, analyte et méthode			
	Analyse et motif de non analyse			
	Méthode d'analyse et la donnée confirmation			
	Plusieurs analyses d'un même RAI		Consignes non appliquées	
	Les résultats et leur interprétation	Paramétrage du Lims		

1.10 Axes d'amélioration

La mise en évidence des facteurs influençant la qualité des données doit aboutir à la proposition d'axes d'amélioration. Ces axes doivent reprendre les facteurs préalablement identifiés par type de facteurs plutôt que par donnée.

Par ailleurs, il est nécessaire de faire des propositions équilibrées et réalistes : les axes d'amélioration doivent être assez ambitieux pour permettre réellement des progrès, mais ils ne doivent pas être utopiques au risque de ne pas être atteints.

Exemple d'évaluation : le dispositif de déclaration des avortements

Matériel

Les feuilles de commémoratifs sont différentes d'un département à l'autre, et de nombreuses données inutiles et/ou inutilisées sont parfois demandées (Annexe II). La collecte est une étape importante pour obtenir des données de bonne qualité : il est donc indispensable de mettre en place une feuille de commémoratifs claire, précise et homogène au niveau national. Notamment, les points suivants sont à prendre en compte :

- sur le fond, seules les données qui seront saisies doivent être demandées :
 - o les données nécessaires à la gestion (non traitées ici) ;
 - o des informations sur l'exploitation : numéro EDE, raison sociale et atelier ;
 - o des informations sur le bovin : numéro IPG complet, date de l'avortement et durée de gestation ;
 - o des informations sur la visite : date, prélèvements réalisés ;
- sur la forme :
 - o les modalités de renseignement de la durée de gestation doivent limiter les erreurs de format et de validité, en proposant par exemple des cases à cocher et en considérant que cette donnée peut être non précisément connue, voire inconnue (moins de 5 mois, 6 mois, 7 mois, 8 mois, 8 mois et demi, à terme, non connue) ;
 - o chaque donnée doit être explicitement demandée et bien différenciée des autres, en particulier entre les dates de visite et d'avortement.

Par ailleurs, sachant que beaucoup de VS renseignent la feuille de commémoratifs après la visite, la possibilité pour eux d'imprimer cette feuille via BDIVet⁹ doit être envisagée : au retour d'une visite pour une DA, le VS pourrait imprimer une feuille de commémoratifs spécifique à la femelle concernée. L'identification de l'exploitation, de l'animal et du vétérinaire seraient pré-renseignées ; la qualité de ces données en serait nettement améliorée.

Comme nous l'avons vu, il y a assez peu de contrôles effectués par SIGAL. La présence de certains blocages serait un atout considérable, en particulier concernant les critères attestant que les DA respectent les spécifications générales d'enregistrement (par exemple, SIGAL ne pourrait autoriser l'export d'un RAI relatif à une DA uniquement s'il ne concerne qu'un bovin ou alors autoriser le renseignement

⁹ BDIVet est une application réservée aux VS qui leur permet notamment d'avoir une interface avec SIGAL.

d'un commémoratif par animal). De plus, il faudrait prévoir la possibilité de renseigner la date d'avortement et la durée de gestation comme « non connue ».

Méthode

De bons outils ne sont utiles et efficaces que si les personnes concernées savent s'en servir. Ils doivent donc impérativement être accompagnés de consignes d'utilisation claires et à tous les niveaux (collecte et saisie des données). Diverses consignes nationales sont donc à envisager :

- un ordre de méthode accompagnant la fiche de commémoratifs nationale, dans le but de s'assurer que les données et les modalités de collecte et de saisie sont bien comprises par tous : définition des données et du format attendu, et définition de la méthode à suivre pour déterminer la durée de gestation (par exemple la taille du fœtus et/ou la date de l'insémination artificielle fécondante) ;
- un ordre de méthode définissant les modalités de validation des données, notamment la gestion des données manquantes ou erronées (par exemple quelles sont les données dont la qualité est jugée prioritaire et dans quels cas le vétérinaire doit être rappelé) ;
- un ordre de méthode définissant les modalités de gestion informatique de la saisie des DA, et précisant en particulier les modalités de gestion des interventions ne respectant pas le schéma relationnel des données et des RAI partiels.

Par ailleurs, nous avons vu que le paramétrage des Lims a une grande marge d'amélioration. Il doit être réalisé dans chaque laboratoire agréé afin d'optimiser au mieux la qualité des données. Pourtant, un certain nombre de consignes existent déjà au sein de la fiche de plan EBRUBVA. Un ordre de méthode pourrait les rappeler en y ajoutant des points complémentaires, notamment :

- rendre obligatoire la présence d'un identifiant bovin au bon format ;
- ne pas rendre obligatoire des données qui peuvent être non connues, comme la date d'avortement et la durée de gestation (tant que ces données ne peuvent être renseignées comme « non connue », ce qui serait l'idéal).

De plus, un contrôle de ce paramétrage, c'est-à-dire la vérification du respect du cahier des charges mis en place par la DDecPP ou par le CoSiR, peut être envisagé.

Enfin, il faudrait envisager une méthode de suivi des indicateurs et d'évaluation régulière du dispositif. Pour cela, des objectifs concrets et réalistes pourraient être mis en place afin d'assurer une amélioration effective de la qualité. A ce titre, deux types d'objectifs peuvent être envisagés, sous la forme d'indicateurs de suivi :

- des indicateurs cibles, identiques pour tous les départements. Ils peuvent s'inspirer de ceux utilisés pour le calcul des IS départementaux. Ils doivent correspondre à ce qui est attendu de la qualité des données du dispositif à long terme ;
- des indicateurs à plus court terme, adaptés à chaque département. Du fait de l'importante disparité interdépartementale mise en évidence et des

fonctionnements différents de chaque département (règles de gestion locales), les évolutions et les améliorations seront beaucoup plus rapides et efficaces si elles sont personnalisées. Ces objectifs peuvent être fixés à l'aide des résultats de l'évaluation menée ici et en concertation avec l'échelon central.

Acteurs

La mise en place de nouvelles consignes devrait être associée à une communication adaptée permettant la promotion de ces consignes.

Tous les acteurs du dispositif peuvent influencer de façon importante la qualité des données. Ils devraient donc être impliqués dans la démarche d'amélioration de la qualité des données, et en premier lieu être sensibilisés à l'importance de disposer de données de bonne qualité. Plus précisément, leur sensibilisation devrait porter sur

- les consignes déjà existantes et non respectées. En particulier, il est primordial de rappeler les analyses à réaliser face à un avortement sachant que plus de la moitié des DA ne respectent pas les consignes de la fiche de plan. Ces écarts ont de lourdes conséquences : un dépistage non fiable et un surcoût pour l'Etat ;
- les nouvelles consignes mises en place, afin que les acteurs en aient connaissance et les appliquent ;
- les objectifs globaux et locaux (voir ci-après).

Cette sensibilisation doit être menée de façon homogène et cohérente, aux niveaux national et départemental. De plus, il est important qu'elle ne soit pas uniquement ponctuelle. Pour maintenir cette sensibilisation dans le temps, l'animation du dispositif joue un rôle important, et devrait également permettre de suivre régulièrement les résultats et d'en discuter avec les acteurs de terrain.

Par ailleurs, le temps consacré à l'application des consignes et notamment à la validation de différentes données devrait être pris en compte. Pour cela, la validation des données devrait apparaître dans les fiches de poste du personnel en charge de cette étape et du temps doit être réservé pour sa réalisation.

ACTIONS A METTRE EN PLACE APRES L'EVALUATION DE LA QUALITE DES DONNEES AU SENS STRICT

La réussite d'une démarche d'évaluation et d'amélioration de la qualité des données dépend fortement et directement de l'implication des acteurs dans les différentes étapes. Cette démarche doit donc être approuvée par tous. De plus, la mise en place de la communication en préalable à la mise en œuvre de l'évaluation de la qualité des données est également importante.

Il en est de même après la mesure de la qualité des données et l'exploitation des résultats. La proposition des axes d'amélioration ne doit en aucun cas marquer la fin du projet :

- le retour d'expérience de l'évaluation et les axes d'amélioration proposés doivent être discutés avec les différents acteurs ;
- les résultats doivent être communiqués aux personnes qui ont participé à l'étude, ainsi qu'à toutes les personnes impliquées dans le dispositif étudié.

Le suivi de l'évaluation est normalement attribué à une personne ou un groupe de personnes dès le lancement de l'évaluation. Cette partie détaille les modalités de la gestion post-évaluation.

1.11 Retour d'information

Le retour d'information doit traiter de deux aspects.

Tout d'abord, il doit concerner la réalisation de l'évaluation en elle-même en soulignant son importance. La communication à ce sujet doit aller dans le sens de celle menée préalablement à l'évaluation, comme un continuum : intérêt de la qualité des données et donc intérêt de cette évaluation.

Ensuite, il doit concerner les résultats de l'évaluation. Les acteurs doivent pouvoir se faire une idée de la qualité des données du dispositif auquel ils participent et des points vers lesquels les efforts à fournir doivent être orientés en priorité.

1.12 Mise en œuvre des axes d'amélioration

L'évaluation a normalement abouti à la proposition d'axes d'amélioration de la qualité des données. Ces axes doivent servir de base à la prise de mesures concrètes, permettant d'améliorer la qualité des données et éventuellement dans un même temps le fonctionnement du dispositif.

Les axes, et les mesures qui en découlent, doivent être discutés avec les différents acteurs et validés par le comité technique, puis par le comité de pilotage du dispositif quand il existe. Il s'agit ainsi de prendre en considération les éventuelles contraintes pratiques et techniques, notamment les moyens humains et matériels disponibles.

1.13 Suivi régulier de la qualité des données

L'évaluation de la qualité des données telle qu'elle est présentée dans ce guide est une action ponctuelle. L'effectivité et l'efficacité de la mise en place d'actions correctives doivent être estimées. Pour cela, un suivi régulier doit être mis en place. Ce suivi peut consister en la mesure des indicateurs construits pour l'évaluation de la qualité des données présentée ici, et la discussion sur l'évolution de la qualité des données. Ce suivi doit permettre, en plus de s'assurer que les mesures mises en place ont entraîné les effets escomptés, d'adapter les axes d'amélioration proposés et éventuellement d'en supprimer et d'en ajouter de nouveaux.

BIBLIOGRAPHIE

Références bibliographiques

- AKOKA, J., BERTI-EQUILLE, L., BOUCELMA, O., BOUZEGHOUB, M., COMYN-WATTIAU, I., COSQUER, M., et al. A framework for quality evaluation in data integration systems. *In: CARDOSO, J., CORDEIRO, J. & FILIPE, J., eds. 9th International Conference on Enterprise Information Systems, June 12-16 2007, Funchal, Madeira - Portugal.*
- BELOEIL, I. (2003). *Analyse de la mise en oeuvre du programme 2002 d'estimation de la prévalence de la tremblante chez les petits ruminants*. Rapport de stage, ENSV.
- BERTI-EQUILLE, L., COMYN-WATTIAU, I., COSQUER, M., KEDAD, Z., NUGIER, S., PERALTA, V. N., et al. 2011. Assessment and analysis of information quality: a multidimensional model and case studies. *IJIQ*, 2, 300-323.
- BRONNER, A., HENAU, V., VERGNE, T., VINARD, J. L., MORIGNAT, E., HENDRIKX, P., et al. 2013. Assessing the mandatory bovine abortion notification system in France using unilist capture-recapture approach. *PLoS One*, 8, e63246.
- CALAVAS, D., FEDIAEVSKY, A., COLLIN, E., TOURATIER, A., AMAR, P., MOQUAY, V., et al. 2012. Plateforme nationale de surveillance épidémiologique en santé animale : missions prioritaires et organisation. *Bulletin épidémiologique Santé animale-alimentation*, 48, 2-5.
- CIHI 2009. The CIHI Data Quality Framework. *In: INFORMATION, C. I. F. H. (ed.)*. Ottawa, Ont.: CIHI.
- DUFOUR, B. & HENDRIKX, P. 2007. *La surveillance épidémiologique en santé animale*, Maisons-Alfort - Paris, Quae - AEEMA.
- GERMAN, R. R., LEE, L., HORAN, J., MILSTEIN, R., PERTOWSKI, C. & WALLER, M. 2001. Updated guidelines for evaluating public health surveillance systems - Recommendations from the guidelines working group. *MMWR*, 50, 1-35.
- HENDRIKX, P., GAY, E., CHAZEL, M., MOUTOU, F., DANAN, C., RICHOMME, C., et al. 2011. OASIS: an assessment tool of epidemiological surveillance systems in animal health and food safety. *Epidemiol Infect*, 139, 1486-96.
- HOINVILLE, L. J., ALBAN, L., DREWE, J. A., GIBBENS, J. C., GUSTAFSON, L., HASLER, B., et al. in press. Proposed terms and concepts for describing and evaluating animal-health surveillance systems.
- JEMM RESEARCH 2008. Des données de qualité : Exploitez le capital de votre organisation. 26.
- KERR, K. & NORRIS, T. 2004. The development of a healthcare data quality framework and strategy. *IQ2004*.
- KERR, K., NORRIS, T. & STOCKDALE, R. Data quality information and decision making: a healthcare case study. 18th Australasian Conference on Information Systems, Toowoomba, 2007. Citeseer, 5-7.
- MAURINO, A. & BATINI, C. 2009. Methodologies for data quality assessment and improvement. *ACM Computing Surveys*, 41, 1-52.
- PALUSSIÈRE, M., CALAVAS, D. & BRONNER, A. 2013. Evaluation de la qualité des données du dispositif de déclaration obligatoire des avortements chez les bovins. *Bull. Epid. Santé Anim. Alim.*, 58, 17-20.
- PIPINO, L. L., LEE, Y. W. & WANG, R. Y. 2002. Data quality assessment. *Communications of the ACM*, 45, 211-218.
- VASSILIADIS, P., BOUZEGHOUB, M. & QUIX, C. 2000. Towards quality-oriented data warehouse usage and evolution. *Information Systems*, 25, 89-115.
- WANG, R. Y. & STRONG, D. M. 1996. Beyond accuracy: What data quality means to data consumers. *J. Manage. Inform. Syst.*, 5-33.

Références réglementaires

- ANONYME 2003. Annexe 1 : Présentation du système d'échange de données informatisées. *Echanges de données informatisées pour les analyses de laboratoire DGAL - Laboratoires agréés Santé animale - Hygiène alimentaire, Cahier des Charges v.1.0.1.*
- ANONYME 2005a. Décision de la Commission du 28 octobre 2005.

ANONYME 2005b. Note de service du 08 novembre 2005 relative à la prophylaxie de la brucellose bovine. Application de l'arrêté du 3 novembre 2005.

ANONYME 2007. Arrêté du 19 décembre 2007 fixant les conditions générales d'agrément des laboratoires d'analyses dans le domaine de la santé publique vétérinaire et de la protection des végétaux. *JO du 19/01/2008*, texte 21.

ANONYME 2008a. Arrêté du 22 avril 2008 fixant les mesures techniques et administratives relatives à la prophylaxie collective et à la police sanitaire de la brucellose des bovinés. *JO du 02/05/2008*, texte 16.

ANONYME 2008b. Note de service du 17 décembre 2008 portant modification de la NS 2008-8242 du 16 septembre 2008 relative à la brucellose des bovinés.

ANONYME 2009a. Arrêté du 22 décembre 2009 relatif à la lutte contre les infections à Salmonella dans les troupeaux de poulets de chair et de dindes d'engraissement, mentionnée à l'article D.223-21, et fixant les modalités de déclaration des salmonelloses aviaires visées à l'article D. 223-1 du code rural. *JO du 29/12/2009*, texte 31.

ANONYME 2009b. Lettre à diffusion limitée DGAL/SDPPST du 06 mai 2009 relative à l'enregistrement des déclarations avortements petits ruminants dans SIGAL.

ANONYME 2010a. Note de service du 24 novembre 2010 portant modification de la note DGAL/SDSPA 2010-8252 relative à la brucellose des bovinés.

ANONYME 2010b. Note de service du 31 août 2010 relative à la brucellose des bovinés : application de l'arrêté du 22 avril 2008 révisé.

ANONYME 2011a. Note de service du 04 octobre 2011 relative à la formation complémentaire du programme national 2011 de formation continue des vétérinaires sanitaires du groupe d'activités n°2.

ANONYME 2011b. Note de service du 14 juin 2011 relative à la délégation de missions administratives relatives à la surveillance sanitaire des exploitations au regard des MRC - convention Etat / GDS - Année N.

ANONYME 2012a. Note de service du 11 septembre 2012 relative au protocole de surveillance de la fièvre Q à mettre en place dans les départements pilotes en lien avec la surveillance de la brucellose.

ANONYME 2012b. Note de service du 12 décembre 2012 relative à la modification de la note de service DGAL/SDSSA/SDPRAT/N2011-8290 relative à la mise en oeuvre du programme de référence SIGAL 07 "Action sanitaire sans les espèces volailles et gibier" - actes de référence "salmonelles aviaires".

ANONYME en vigueur du 7 août 2003 au 2 juillet 2012. Article R*223-79 du Code rural.

Site internet

ANSES. (Dernière consultation le 26 juillet 2013). *Site du Bulletin Epidémiologique Santé animale-alimentation - Anses, DGAL*. Adresse URL: <http://www.ansespro.fr/bulletin-epidemiologique/index.htm>.

Annexe I – Formulaire d'évaluation de la qualité des données

Section 1 - Description du dispositif et des modalités de gestion des données

A – Cadrage de l'évaluation

Objet de l'évaluation :

Objectif principal de l'évaluation :

Personnes concernées par l'évaluation :

-
-
-

B – Description des modalités de mise en œuvre du dispositif

Objet de la surveillance :

Objectifs du dispositif :

-
-
-
-

Animation et maîtrise d'ouvrage :

-
-
-

Modalités de la surveillance :

- Surveillance événementielle ou active :
- Population surveillée :
- Echantillonnage :

Modalités de réalisation de l'acte :

- Réalisation de la visite :
- Prélèvements réalisés :

Modalités d'analyse :

Modalités de gestion des données :

Modalités de formation des acteurs :

Modalités de communication :

Modalités d'évaluation du dispositif :

Entité élémentaire de surveillance :

Etude par animal ou par groupe d'animaux :

Documents disponibles :

Références réglementaires

-
-
-
-

Références non réglementaires

-
-

-
-

C – Description du flux des données

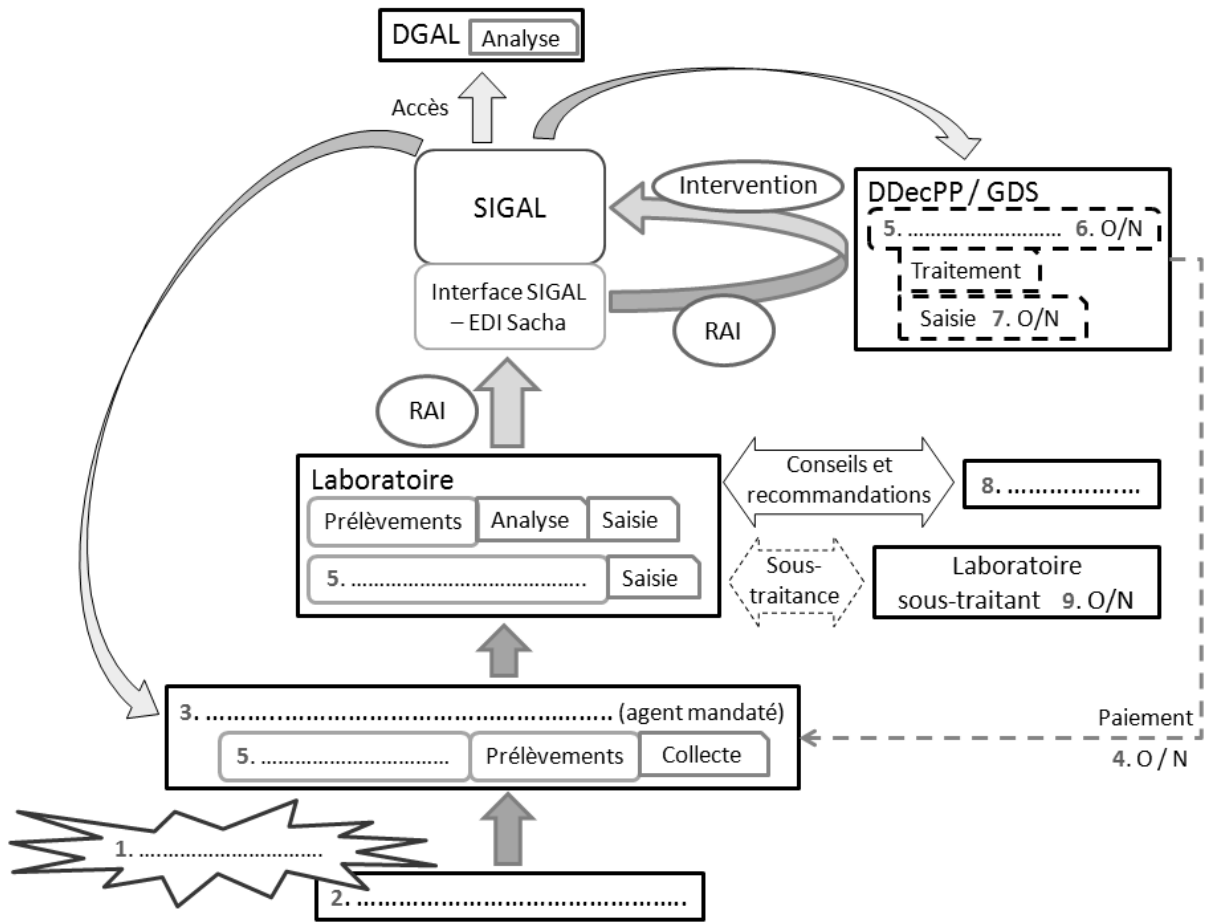


Figure An.1 : Schéma de circulation des données pour une gestion par INP à remplir

1. Evènement surveillé
2. Personne déclarant l'évènement
3. Personne réalisant les prélèvements
4. Rémunération par l'Etat ou non de la personne réalisant les prélèvements
5. Support de la collecte des données
6. Transmission ou non du support à la DDecPP ou au GDS
7. Saisie ou non de données dans SIGAL par la DDecPP ou le GDS
8. LNR (ou équivalent)
9. Possible sous-traitance ou non des analyses

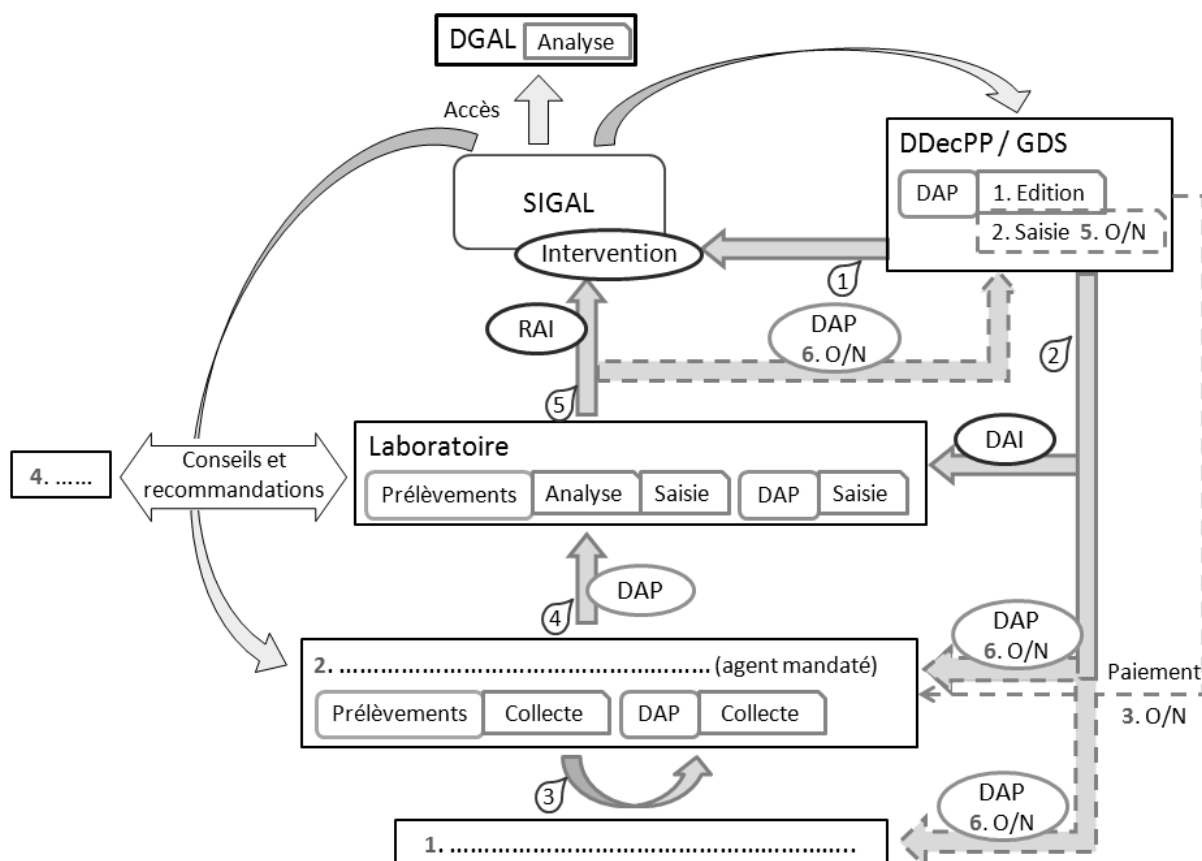


Figure An.2 : Schéma de circulation des données pour une gestion par IP à remplir

1. Personne responsable des animaux faisant l'objet de l'intervention
2. Personne réalisant les prélèvements
3. Rémunération par l'Etat ou non de la personne réalisant les prélèvements
4. LNR (ou équivalent)
5. Saisie ou non des données dans SIGAL par la DDecPP ou le GDS
6. Circuit de transmission du DAP

D – Réalisation d'un schéma relationnel des données

Dans le tableau An.1, entourer les contraintes souhaitées.

Tableau An.1 : Contraintes souhaitées pour le SRD

Type de relation	Contraintes imposées	Contraintes souhaitées
1 : nombre d'interventions par EES	0,n
2 : nombre d'EES par intervention	1,n
3 : nombre d'EES par RAI	1,n
4 : nombre de RAI par intervention	0,n
5 : nombre de RAI par EES	0,n

Section 2 – Sélection et description des données

→ Sélection

Liste des données retenues

- Commémoratifs :
 -
 -
 -
 -
- Prélèvement
- Matrice
- Analyte
- Méthode
- Résultat
- Interprétation
- Identifiant de la structure rattachée à l'intervention
- Identifiant de l'animal ou du groupe d'animaux
-
-
-
-

→ Description

Le tableau An.2 est à remplir pour chaque donnée sélectionnée.

Tableau An.2 : Description des données à l'aide de métadonnées

Nom	
Définition	
Obligation	
Format	
Valeurs possibles	
Mise à jour	
Source	

→ Catégorisation

Distribuer les différentes données analysées dans les cases correspondantes :

Tableau An.3 : Catégorisation des données en fonction du niveau de renseignement et du degré de structuration

	Données structurées	Données semi-structurées	Données non structurées
Renseignement obligatoire			
Renseignement non obligatoire			

Section 3 – Evaluation qualitative de la qualité des données

Tableau à remplir par la personne menant l'entretien :

Tableau An.4 : Notation de la compréhension

Nom de la donnée	Compréhension
	<i>Note /3</i>

Tableau à remplir par la personne interrogée :

Tableau An.5 : Notation de la pertinence, de la crédibilité et de l'accès et utilisation

Nom de la donnée	Collecte	Saisie	Accès	Pertinence (si accès)	Crédibilité (si collecte ou saisie)	Accès et utilisation (si accès)
	<i>Oui ou non</i>	<i>Oui ou non</i>	<i>Oui ou non</i>	<i>Note /3</i>	<i>Note /3</i>	<i>Note /3</i>

Tableau à remplir par la personne menant l'entretien :

Tableau An.6 : Commentaire sur l'évaluation qualitative de la qualité des données

Nom de la donnée	Pertinence	Accès et utilisation	Crédibilité	Divers
	<i>A quoi la donnée est-elle utile</i>	<i>Est-elle utilisée ? A quelle fréquence ? Pour quoi ?</i>	<i>En quoi la donnée n'est pas crédible ? Est-il question d'objectivité ?</i>	

Section 4 – Evaluation des spécifications générales d'enregistrement

Compléter le tableau An.7 :

Tableau An.7 : Critères et indicateurs relatifs au respect des spécifications générales d'enregistrement

Critère	Indicateur
Validité de l'identifiant de la structure où est rattachée l'intervention	Nombre d'interventions qui respectent le critère / Nombre total d'interventions
Les interventions en doublon	Nombre d'entités élémentaires de surveillance qui respectent le critère / Nombre total d'entités
Date visite antérieure à la date RAI (de tous les RAI s'il y en a plusieurs)	Nombre d'interventions qui respectent le critère / Nombre total d'interventions
.....

Section 5 - Evaluation quantitative de la qualité des données respectant les spécifications générales d'enregistrement

A – Evaluation à l'échelle des données

→ Complétude, format et validité

Pour chaque dimension, lister les données retenues et indicateurs correspondant.

Complétude :

Tableau An.8 : Données pour lesquelles la complétude est évaluée et indicateurs correspondant

Nom de la donnée	Indicateur
	<i>Du type : nombre de données renseignées / nombre de données dont le renseignement est attendu</i>

Format :

Tableau An.9 : Données pour lesquelles le format est évalué et indicateurs correspondant

Nom de la donnée	Indicateur
	<i>Du type : nombre de données au bon format / nombre de données renseignées</i>

Validité :

Tableau An. 10 : Données pour lesquelles la validité est évaluée et indicateurs correspondant

Nom de la donnée	Indicateur
	<i>Du type : nombre de données valides / nombre de données renseignées et au bon format</i>

B – Evaluation à l'échelle de l'entité élémentaire de surveillance

Tableau An.11 : Dimensions étudiées pour l'évaluation à l'échelle de l'EES et indicateurs correspondant

Dimensions		Indicateur
Vitesse de circulation des données	<i>Liste des délais</i>	<i>De type : délai moyen en jours</i>
Complétude		Nombre d'entités pour lesquelles toutes les données sont renseignées / nb total d'entités
Validité		Nombre d'entités pour lesquelles toutes les données sont valides / nb d'entités où elles sont toutes renseignées

C – Evaluation de la cohérence

Tableau An.12 : Cohérence étudiées et indicateurs correspondant

Cohérences étudiées	Indicateur
1. Prélèvement - matrice analysée	Nombre de couples de données respectant la cohérence étudiée / nombre de couples de données total
2. Matrice analysée - analyte recherché - méthode d'analyse (MAM)	
3. MAM - motif de non analyse	
4. MAM - analyse de confirmation	
5. Différents MAM d'un RAI	
6. Résultats – interprétation	
Autres	

Section 6 – Indicateurs synthétiques

Tableau An.13 : Valeurs cibles des indicateurs spécifiques, regroupés par indicateur synthétique

Indicateur synthétique	Indicateur spécifiques	Valeurs cibles proposées
IS 1	<i>Nom de l'indicateur</i>	<i>En pourcentage</i>
IS 2		

Section 7 – Recherche de corrélations

Tableau An.14 : Corrélation étudiées et résultats du test de Pearson

Corrélations	P-value	Coefficient de corrélation (si p-value < 0,05)

Annexe III – Description des données à l'aide de métadonnées

La description des données suit la logique du SRD : les données sont regroupées par entité du SRD. Le support pris en compte par défaut est l'intervention. Lorsqu'une métadonnée est différente en fonction du support, sa valeur pour la donnée dans le RAI est précisée entre parenthèses. Certaines interrogations n'ont pas pu être résolues.

Les métadonnées utilisées sont les suivantes :

- nom de la donnée ;
- définition de la donnée : ce qu'elle représente ;
- obligation : indique si la présence de l'élément est obligatoire ou optionnelle ;
- format de la donnée : format sous lesquelles la donnée peut être renseignée ;
- valeurs possibles : les différentes valeurs que peut prendre la donnée ;
- mise à jour : la possibilité ou non et les modalités d'une mise à jour de la donnée ;
- source : d'où provient la donnée.

a- Les données relatives à la visite du vétérinaire

Nom	Date de visite	Numéro d'ordre
Définition	Date de la visite du VS	Numéro d'identification du vétérinaire sanitaire ou de son cabinet
Obligation	Oui	Oui
Format	Date, contraint (mais pas dans les lims ?)	Alphanumérique, contraint
Valeurs possibles	Absence de contrainte	ORDRE + un nombre de 3 à 6 chiffres, ou CSOPRO + un nombre à 6 chiffres (<500 000, c'est un numéro personnel, > 500 000 c'est un numéro de cabinet), liste préenregistrée
Mise à jour	?	Le VS mentionné dans le RAI peut être modifié dans SIGAI
Source	Commémoratifs	Commémoratifs

Nom	Type d'identifiant de l'exploitation	Valeur de l'identifiant	Atelier
Définition	Base de référence à laquelle correspond l'identifiant	Identifiant de l'exploitation	Identifiant de l'atelier, type de production, type d'atelier
Obligation	Oui	Oui	Non (que si le numéro d'identification est distinct de celui de l'exploitation)
Format	Alphanumérique, contraint	Numérique, contraint	Alphanumérique, contraint (mais pas dans les Lims)
Valeurs possibles	EDE, SIRET, SIREN, NAGRI, ILU, RESA, NPAC, DOSDSV, CGIDIC, NAPI, PROJET, ELEVAG, DEPADM, ORDRE, TVA	Type EDE : nombre à 8 chiffres (département / commune / établissement), liste déroulante	Type EDE : 8 chiffres (dpt/ commune/établissement) + production + type d'atelier
Mise à jour	Possibilité de supprimer l'intervention rattachée à l'établissement		Possible
Source	Commémoratifs	Commémoratifs	Commémoratifs

b- Les données relatives à la femelle prélevée

Nom	Date d'avortement	Durée de gestation	Numéro du bovin
Définition	Date à laquelle la femelle a avorté	Stade de gestation au moment de l'avortement	Numéro national d'identification de la femelle
Obligation	Non	Non	Non
Format	Date, contraint (mais pas dans les Lims)	Numérique en mois, contraint (mais pas dans les Lims)	Alphanumérique, non contraint
Valeurs possibles	Absence de contrainte	0 à 12 (contraint uniquement dans SIGAI)	FR suivi de 10 chiffres (non contraint)
Mise à jour	Possible	Possible	Impossible
Source	Commémoratifs	Commémoratifs	Commémoratifs

c- Les données relatives aux prélèvements

Nom	Matrice prélevée	Motif non analyse
Définition	Type de prélèvement réalisé	Justification de la non-réalisation d'analyse à partir de la matrice prélevée
Obligation	Oui	Non
Format	Alphabétique, contraint	Alphabétique, contraint
Valeurs possibles	Sang ou écouvillon vaginal (<i>placenta, avorton, mucus vaginal, nœud lymphatique, lait individuel</i>)	Analyse conditionnelle, contamination, contenant vide, contenant cassé, échantillon non conforme, échantillon corrompu, sérum hémolysé, quantité insuffisante, autre
Mise à jour	Impossible	Impossible
Source	Commémoratifs	Laboratoire

d- Les données relatives au plan d'analyse

Nom	Nom du plan	Plan complet
Définition	Identification du plan d'analyse mis en œuvre	Renseigne si le plan est complet ou non
Obligation	Oui	Oui
Format	Alphabétique, contraint	Alphabétique, contraint
Valeurs possibles	Elevage-Brucellose bovine-Avortement	O (oui) ou N (non)
Mise à jour	Impossible	Impossible
Source	Laboratoire	Laboratoire

e- Les données relatives aux analyses

Nom	Matrice analysée	Analyte	Méthode	Confirmation
Définition	Constituant de l'échantillon analysé	Objet de la méthode d'analyse	Méthode d'analyse utilisée pour détecter l'analyte	Renseigne sur le but de l'analyse : confirmation ou dépistage
Obligation	Non	Non	Non	Non
Format	Alphabétique, contraint	Alphabétique, contraint	Alphabétique, contraint	Alphabétique, contraint
Valeurs possibles	Sérum ou écouvillon vaginal (<i>mucus, avorton, placenta, nœud lymphatique, lait</i>)	Brucellose bovine-Anticorps	EAT, ELISA, FC, culture, identification	O (oui) ou N (non)
Mise à jour	En cas d'erreur, un nouvel RAI doit être envoyé, une nouvelle intervention créée (et la précédente supprimée)			
Source	Laboratoire-commémoratifs	Laboratoire	Laboratoire	Laboratoire

Nom	Résultat	Date de validation
Définition	Résultat d'un MAM ¹⁰	Date à laquelle le laboratoire valide le résultat du RAI
Obligation	Non	Oui
Format	Alphabétique, contraint	Date, contraint
Valeurs possibles	POS (positif) ou NEG (négatif)	Pas de contrainte
Mise à jour	En cas d'erreur, un nouvel RAI doit être envoyé, une nouvelle intervention créée	
Source	Laboratoire	Laboratoire

f- Les données relatives au RAI

Nom	Nom du RAI	Date du RAI	Etat du RAI
Définition	Identifiant du résultat d'analyses informatisé	Date à laquelle le RAI a été reçu dans SIGAI	Renseigne si le RAI est complet ou non
Obligation	Oui	Oui	Oui
Format	Alphanumérique, contr	Date, contraint	Alphanumérique, contraint
Valeurs possibles	RA01 + laboratoire + DDPP/GDS + département + 15 chiffres	Pas de contrainte	C (complet), P (partiel), F (fin d'un RAI partiel)
Mise à jour	En cas d'erreur, un nouvel RAI doit être envoyé, une nouvelle intervention créée		
Source	RAI (automatique)	SIGAI (automatique)	Laboratoire

Nom	Laboratoire	Laboratoire sous-traitant	Numéro de lims
Définition	Laboratoire qui reçoit les échantillons et envoie le RAI	Laboratoire auquel une analyse a été sous-traitée (ou plusieurs)	Identifiant informatique du laboratoire
Obligation	Oui	Non	Oui
Format	Alphanumérique, contraint	Alphanumérique, contraint	Fonction du Lims
Valeurs possibles	LVD + numéro du dép.	Tout laboratoire accrédité	Fonction du Lims
Mise à jour	En cas d'erreur, un nouvel RAI doit être envoyé, une nouvelle intervention créée		
Source	Laboratoire	Laboratoire	RAI (automatique)

¹⁰ Un MAM est un triplet matrice analysée, analyte, méthode.

Nom	Interprétation du laboratoire
Définition	Interprétation du ou des résultats d'un plan d'analyse par le laboratoire
Obligation	Non
Format	Alphabétique, contraint
Valeurs possibles	CONFORM, ENCOURS, IN_ITPBLE, NCONF, NEG, POS
Mise à jour	En cas d'erreur, un nouvel RAI doit être envoyé, une nouvelle intervention créée
Source	RAI

g- Les données relatives à l'intervention

Nom	Numéro de l'intervention	Type de l'intervention	Synthèse
Définition	Identifiant de l'intervention	Nature de l'intervention réalisée	Synthèse globale des résultats de l'intervention
Obligation	Oui	Oui	Non
Format	Nombre réel, contraint	Alphabétique, contraint	Alphanumérique, contraint
Valeurs possibles	1 ou 2 (DDPP/GDS) + département + 8 chiffres	Enregistrement d'une déclaration d'avortement bovin	Niveau 1 : vert, Niveau 2 : orange, Niveau 3 : rouge
Mise à jour	En cas d'erreur l'intervention doit être supprimée		Possible
Source	SIGAI (automatique)	DDPP/GDS	SIGAI (automatique)

Nom	Département	Date	Motif de non paiement
Définition	Département de création de l'intervention	Date de création de l'intervention	Raison d'un non paiement du VS
Obligation	Oui	Oui	Non
Format	Nombre réel, contraint	Date, contraint	Alphabétique, contraint
Valeurs possibles	0 à 96	Pas de contrainte	AUT – autre ou DBL – doublon
Mise à jour	En cas d'erreur l'intervention doit être supprimée		Possible
Source	SIGAI (automatique)	SIGAI (automatique)	DDPP / GDS

Annexe IV – Document support pour l'évaluation qualitative de la qualité des données

1- Présentation de la démarche globale d'évaluation de la qualité des données

Titre de l'étude

La qualité des données collectées dans le cadre d'un dispositif de surveillance en santé animale est un enjeu central. En effet, la qualité des données influence directement l'évaluation du fonctionnement du dispositif à l'aide d'indicateurs de fonctionnement, et de la capacité du dispositif à répondre à ses objectifs (par exemple le délai de dépistage conditionne la capacité de détection précoce). Elle influence également la qualité de l'évaluation de la situation sanitaire qui se fait par l'analyse des données collectées. Ainsi, la qualité de ces données, et notamment de leur capacité à aboutir à une information fiable, c'est-à-dire à la fois juste (conforme à ce qui est réellement observé sur le terrain) et pouvant être traitée sans erreur (par exemple, la donnée est conforme au format attendu), influence indirectement les décisions des pouvoirs publics (qui reposent sur les indicateurs de fonctionnement du dispositif et sur l'évaluation de la situation sanitaire). Son évaluation apparaît donc être un préalable indispensable.

L'objectif de cette étude... (mentionner également les personnes encadrant l'étude).

Objectif principal du déplacement :

Points à discuter :

Contact

2- Présentation de la phase qualitative de l'évaluation de la qualité des données

Evaluation qualitative de la qualité des données du dispositif de ...

Une donnée est définie comme un fait, une valeur affectée à une variable (les notions de donnée et de variable seront ici indifféremment associées à l'appellation « donnée »). Elle se distingue de l'information qui correspond à une donnée traitée de façon à améliorer la connaissance de la personne qui utilise la donnée.

Une donnée de qualité élevée correspond à une donnée qui *est convenable pour l'usage que l'on souhaite en faire* (« *fitness for use* »). Ainsi, cette notion est étroitement dépendante des utilisateurs de la donnée et des personnes conduisant l'évaluation de la qualité des données.

La qualité d'une donnée est un concept multidimensionnel, qui se définit comme l'association de plusieurs dimensions. Deux niveaux de dimensions peuvent être définis : un premier niveau qui

correspond aux objectifs globaux de qualité des données habituellement rencontrés, un second niveau constituant des facteurs pouvant influencer les dimensions de premier niveau et dépendant strictement de la perception qu'ont les utilisateurs des données.

Les dimensions de second niveau étudiées sont listées et définies dans le tableau suivant.

Nom	Définition
Pertinence	Utilité de la donnée, qui concorde avec les besoins des utilisateurs
Crédibilité	Perception de l'exactitude de la donnée par l'utilisateur
Accès et utilisation	Facilité d'accès et d'utilisation à un niveau pratique (rapide, sans intervention de personnes extérieures)

Les indicateurs permettant de mesurer ces dimensions sont des notes attribuées par les acteurs du dispositif. La méthode à suivre pour la notation est la suivante :

Dimension	Question	3/3	2/3	1/3	0/3
Pertinence	Pensez-vous que cette donnée soit utile, que ce soit pour la surveillance ou pour vous (gestion, communication...) ?	Oui, elle est indispensable	Plutôt oui	Pas vraiment	Non, pas du tout
Crédibilité	Quelle confiance avez-vous dans l'exactitude (et la précision) de cette donnée ?	Oui, totalement	Oui, assez	Pas vraiment	Non, pas du tout
Accès et utilisation	Avez-vous facilement accès à cette donnée ? Est-ce facile de l'utiliser ?	Très facile	Plutôt facile	Assez compliqué	Très compliqué

Questions associées :

Si la donnée est utile, pour quoi l'est-elle ?

Est-ce que la donnée est utilisée ? Si oui, à quelle fréquence ?