

IUMSP
Institut universitaire de médecine sociale et préventive

Unité d'épidémiologie du cancer - UEC

PERFORMANCES RADIOLOGIQUES DU PROGRAMME VALAISAN DE DEPISTAGE DU CANCER DU SEIN: EVOLUTION ET DETERMINANTS, 2002-12

Jean-Luc Bulliard, Fabio Levi

Raisons de santé 214 – Lausanne 2013

Etude financée par :

Promotion Santé Valais pour le Programme valaisan de dépistage du cancer du sein

Citation suggérée :

Bulliard J-L, Levi F. Performances radiologiques du programme valaisan de dépistage du cancer du sein : évolution et déterminants, 2002-12. Lausanne: Institut universitaire de médecine sociale et préventive, 2013 (Raisons de santé, no 214).

Remerciements :

Dr E. Dayer (Président du Groupe de référence pour le Programme valaisan de dépistage du cancer du sein), M. J-B. Moix (Directeur Promotion Santé Valais), Drs B. Monnier (Radiologue conseil) et B. Filliez (Médecin responsable), et Mme N. Tridondane (Administratrice du Programme valaisan de dépistage du cancer du sein) pour leur étroite collaboration; Dr I. Konzelmann (responsable du Registre valaisan des tumeurs) pour la réalisation d'appariements informatiques et la mise à disposition de données.

Date d'édition :

juillet 2013

Nom du dossier : M:\IMS\UEC\SECRETAR\RAISONS DE SANTE\RdS_214.docx

TABLE DES MATIERES

| | |
|--|-----------|
| Liste des Tableaux..... | 4 |
| Liste des Figures..... | 5 |
| Liste des sigles et abréviations | 6 |
| Résumé | 7 |
| 1 Introduction..... | 9 |
| 1.1 Preambule..... | 9 |
| 1.2 Objectifs | 9 |
| 1.3 Structure et contenu..... | 10 |
| 2 Evolution de l'activité du Programme, 1999-2012 | 11 |
| 2.1 Participantes éligibles | 11 |
| 2.2 Réalisations et lectures des mammographies, 1999-2012..... | 12 |
| 2.3 Actions prises par le programme | 14 |
| 2.4 Commentaires et choix analytiques..... | 14 |
| 3 Evolution des performances du Programme | 16 |
| 3.1 Evolution du taux de reconvocation | 16 |
| 3.2 Analyse des tendances des indicateurs de qualité..... | 18 |
| 3.3 Commentaires..... | 19 |
| 4 Déterminants de la qualité des lectures..... | 20 |
| 4.1 Considérations préliminaires..... | 20 |
| 4.2 Résultats des lectures multiples | 20 |
| 4.3 Déterminants des interprétations radiologiques | 22 |
| 4.4 Commentaires..... | 23 |
| 5 Cancers d'intervalle | 25 |
| 5.1 Identification | 25 |
| 5.2 Fréquence | 26 |
| 5.3 Profil clinico-épidémiologique | 27 |
| 5.4 Commentaires..... | 30 |
| 5.5 Recommandations | 31 |
| 6 Conclusions et recommandations..... | 33 |
| 6.1 Synthèse | 33 |
| 6.2 Recommandations | 34 |
| 7 Annexes | 36 |
| 7.1 Glossaire | 36 |
| 7.2 Mesures d'amélioration prises | 38 |
| 7.3 Résultats des analyses statistiques de tendances des principaux indicateurs de qualité..... | 39 |

LISTE DES TABLEAUX

| | | |
|------------------|---|----|
| Tableau 1 | Statistiques annuelles d'activité du Programme valaisan de dépistage, 1999 - 2012 | 12 |
| Tableau 2 | Valeurs (globales et récentes) et évolution des principaux indicateurs de qualité du Programme valaisan de dépistage | 17 |
| Tableau 3 | Taux de faux positif, de détection, de cancers non vus, et valeur prédictive positive (VPP) selon le contexte de lecture, le type de lecteur et le volume individuel de lecture, tous tours confondus (2002-12) | 23 |
| Tableau 4 | Fréquence des cancers d'intervalle (1999-2010) du Programme valaisan de dépistage par mammographie et comparaison aux normes européennes. | 26 |
| Tableau 5 | Sensibilité et spécificité par type de tour du Programme valaisan de dépistage par mammographie..... | 27 |
| Tableau 6 | Profil clinique des cancers dépistés et d'intervalle du Programme valaisan, 1999-2010..... | 28 |
| Tableau 7 | Profil épidémiologique des cancers dépistés et d'intervalle du Programme valaisan, 1999-2010..... | 29 |
| Tableau 8 | Echange d'informations entre le CVD et le RVsT : éléments perfectibles et mesures proposées..... | 31 |
| Tableau 9 | Description des actions effectuées auprès des radiologues pour améliorer la qualité du Programme valaisan | 38 |

LISTE DES FIGURES

| | | |
|------------------|--|----|
| Figure 1 | Nombre de participations par femme au Programme valaisan de dépistage (octobre 1999 à novembre 2012) | 11 |
| Figure 2 | Evolution du nombre de mammographies et de lecteurs (premiers et deuxièmes) du Programme valaisan de dépistage (10/1999 - 09/2012) | 13 |
| Figure 3 | Evolution du nombre annuel de mammographies par type de lecteurs dans le Programme valaisan de dépistage, 2000 – 2012..... | 13 |
| Figure 4 | Evolution entre 2002 et 2011 du taux de détection et de faux positif du Programme valaisan de dépistage par mammographie en tour prévalent et incident | 16 |
| Figure 5 | Taux de reconvoication et VPP par année et type de tour de dépistage, 2002-12..... | 18 |
| Figure 6 | Résultats des lectures multiples du Programme valaisan en tour prévalent, 2002-2012..... | 21 |
| Figure 7 | Résultats des lectures multiples du Programme valaisan en tour incident, 2002-2012..... | 21 |
| Figure 8 | Flux et classification des cancers chez les participantes (1999-2008) identifiés par le RVsT..... | 25 |
| Figure 9 | Tendances annuelles (2002-11) des principaux indicateurs de qualité du Programme valaisan par tour de dépistage (analyse Joinpoint)..... | 39 |
| Figure 10 | Tendances annuelles de la valeur prédictive positive en 1 ^e et 2 ^e lecture du Programme valaisan (analyse Joinpoint) | 41 |

LISTE DES SIGLES ET ABRÉVIATIONS

| | |
|----------|--|
| BIRADS : | Breast Imaging-Reporting And Data System |
| CVD : | Centre valaisan de dépistage du cancer du sein |
| ICSN : | International Cancer Screening Network |
| IUMSP : | Institut universitaire de médecine sociale et préventive de Lausanne |
| L1 : | 1 ^e lecture |
| L2 : | 2 ^e lecture |
| L3 : | 3 ^e lecture |
| R1 : | Radiologue 1 ^e lecteur |
| R2 : | Radiologue 2 ^e lecteur |
| RVsT | Registre valaisan des tumeurs |
| VPP : | Valeur prédictive positive |
| VPN : | Valeur prédictive négative |
| UEC : | Unité d'épidémiologie du cancer |

RESUME

Lancé en octobre 1999, le Programme valaisan de dépistage du cancer du sein a pour mission de promouvoir, d'organiser et de mener une action de dépistage auprès de la population féminine valaisanne âgée de 50 à 70 ans. Ce programme est régulièrement évalué, de manière indépendante, afin de s'assurer que sa qualité et son efficacité répondent à des normes internationalement admises tout en minimisant ses effets adverses. Ces évaluations ont mis en évidence un niveau de performances globalement satisfaisant et suggéré un potentiel d'amélioration, notamment de la qualité des lectures radiologiques. Dans un souci d'optimiser les performances, la direction du Programme a mené ces dernières années plusieurs actions auprès de ses radiologues et pris quelques mesures.

Ce rapport thématique traite de l'évolution des performances du programme, principalement au cours de la dernière décennie (2002-11), c'est-à-dire après sa phase de démarrage (1999-2001). Les tendances des principaux indicateurs de performance et leurs associations avec les actions et mesures réalisées, qui ont été répertoriées et classées selon leur types et objectifs, sont analysées. Les facteurs les plus influents sur la qualité des lectures radiologiques sont quantifiés de manière à faciliter toute décision interventionnelle. Les cancers d'intervalle (1999-2010) du Programme sont étudiés afin d'en apprécier la fréquence, d'en établir le profil pronostique et de déterminer les caractéristiques clinico-épidémiologiques qui les distinguent des cancers dépistés dans le cadre du Programme.

Entre 1999 et 2012, **l'activité de dépistage** a crû d'environ 5% par an (moyenne annuelle: 9629 mammographies) et plus de 125'000 mammographies ont été réalisées, dont 3 sur 4 concernaient des femmes qui avaient déjà participé au Programme. Le cap des 10'000 mammographies a été franchi en 2007. Au niveau des radiologues, la progression est régulière et plus prononcée pour les 2^e lecteurs dont le volume annuel de lectures a doublé en 10 ans pour atteindre 1500 en 2012 (moyenne annuelle: 1266). Le nombre moyen de mammographies interprétées par les 1^e lecteurs a peu augmenté jusqu'en 2007, s'est accru dès 2008 et a atteint 500 lectures en 2012 (moyenne annuelle: 447). **L'application dès 2007-8 d'une politique de contrôle prospectif plus strict des accréditations, surtout comme 2^e lecteur, a permis dans un contexte de croissance d'activité, d'augmenter, tout en les distribuant de manière plus homogène, les volumes individuels de lecture.**

Sur le plan de l'évolution des performances, les tendances indiquent une amélioration continue, notamment des indicateurs de qualité du dépistage, avec un niveau de performance global conforme aux normes européennes: les taux de reconvocation (et de faux positif) baissent, et la part de cancers invasifs dépistés de petite taille augmente au détriment de celle des cancers de stade avancé. Si la plupart des tendances restent modestes et ne peuvent entièrement exclure comme explication des fluctuations aléatoires, les progrès les plus manifestes concernent les quelques indicateurs pour lesquels le potentiel d'amélioration était le plus grand. Les performances les plus élevées ont systématiquement été mesurées en 2006 et coïncident avec les mesures de sensibilisation personnalisée entreprises alors par la direction du Programme, dont un entretien avec chaque radiologue accrédité.

Le type de tour (première ou multiple participation), le contexte de lectures (1^e, dans un centre de radiologie, ou 2^e, au Centre Valaisan de Dépistage (CVD)) et l'année de réalisation de la mammographie sont les déterminants les plus importants de la qualité des lectures du Programme. **Les radiologues valaisans sont les plus performants lorsqu'ils interprètent des mammographies récentes, au CVD (2^e lecture) et avec des clichés comparatifs (tour incident). En revanche, que les clichés soient analogiques ou numériques ou que la 1^e lecture soit effectuée par un radiologue 1^e ou 2^e lecteur n'influencent que de façon marginale les performances interprétatives.** Le différentiel de volume de lectures encore peu marqué entre 1^e et 2^e lecteurs pourrait expliquer à la fois la faible différence de performances entre les 2 types de lecteur dans un

même contexte ainsi que l'absence d'une association claire entre le volume (annuel ou cumulé) et la qualité des lectures.

La fréquence des **cancers d'intervalle** du Programme, de l'ordre de 2 cas pour 1000 femmes dépistées, est en légère baisse. Elle est en adéquation avec le référentiel européen (25% la première année (norme : <30%), 52% la seconde année (norme : < 50%)) et comparable aux observations faites dans d'autres programmes. Six cancers d'intervalle sur 7 sont survenus après des mammographies négativées par tous les lecteurs et un sur 13 (7,6%) correspond à un faux négatif des investigations complémentaires (mammographies positives du Programme).

Dans le cadre de ce rapport, le profil clinico-épidémiologique des cancers d'intervalle d'un Programme a été évalué pour la première fois en Suisse. Ce profil met en évidence une prédominance de femmes quinquagénaires, à seins denses et sous traitement hormono-substitutif. Surtout, les cancers d'intervalle présentent un profil pronostique clairement plus défavorables que les cancers dépistés. Ces résultats sont conformes aux connaissances épidémiologiques. Le suivi pendant au moins 2 ans de toutes les femmes dont le dépistage s'est révélé négatif, grâce à la collaboration du Registre Valaisan des Tumeurs (RVsT), a permis de vérifier que la sensibilité (71%) et la spécificité (96%) du Programme était satisfaisante.

Plusieurs pistes pour consolider le niveau de performances atteint et poursuivre l'amélioration en cours de la qualité des lectures sont proposées. Quelques mesures pratiques en vue d'améliorer l'échange d'informations entre le CVD et le RVsT pour les cancers d'intervalle et les cancers dépistés sont suggérées. Une revue des cancers d'intervalle, prioritairement dans le cadre d'une étude scientifique, est recommandée afin de mieux cerner les éléments iconographiques qui seraient associés aux cancers d'intervalle et d'y sensibiliser les radiologues. Une accentuation de la réduction du nombre de 2^e lecteurs, afin d'augmenter les volumes annuels de lecture, et l'instauration de manière durable des mesures personnalisées de 2006 dont le succès a pu être démontré, sont fortement conseillées.

En conclusion, ce rapport thématique confirme l'évolution favorable des performances du programme valaisan. Les résultats portent à croire que les mesures de sensibilisation par entretien personnalisé avec chaque radiologue en 2006 et l'adoption en 2007-8 d'une politique plus stricte d'accréditation des lecteurs, surtout pour la 2^e lecture, ont fortement contribué à cette tendance, même si un lien causal entre les actions entreprises et les effets observés ne peut formellement être établi. Ces mesures d'amélioration s'inscrivant dans la durée, les prochaines évaluations devraient confirmer la poursuite des tendances amorcées.

1 INTRODUCTION

1.1 PREAMBULE

Une évaluation périodique des performances d'un programme organisé de dépistage, selon des normes codifiées et reconnues^a, permet de s'assurer de sa qualité et de son efficacité tout en minimisant ses effets adverses. Ces évaluations sont obligatoires en Suisse^b.

L'évaluation du programme valaisan de dépistage du cancer du sein est confiée depuis son démarrage à l'Unité d'épidémiologie du cancer (UEC) de l'Institut universitaire de médecine sociale et préventive (IUMSP) de Lausanne. Cette Unité a acquis une grande expérience dans l'évaluation des actions de prévention contre le cancer, notamment du dépistage par mammographie. L'UEC a par exemple réalisé une dizaine d'évaluations épidémiologiques de programmes romands^c et, plus récemment, conduit pour la Fondation *swiss cancer screening* le premier monitoring national des programmes régionaux de dépistage du cancer du sein^d.

Les évaluations précédentes et le monitoring national ont relevé pour le programme valaisan une participation régulière des femmes, la plus élevée parmi les programmes romands, ainsi qu'un niveau de performances globalement satisfaisant et proche des normes internationalement admises. Ces évaluations ont aussi suggéré un potentiel d'amélioration de la qualité des lectures radiologiques^e. Soucieuse d'optimiser les performances de son programme, la direction du Programme valaisan a mené ces dernières années des actions auprès des radiologues agréés et pris quelques mesures afin d'améliorer la qualité des lectures.

Ce rapport thématique s'intéresse à l'évolution des performances du programme valaisan, principalement sur le plan de la qualité des lectures et de ses déterminants. A cet égard, une exploration détaillée des cancers d'intervalle du programme, après 13 ans (1999-2012) d'activité de dépistage, y est présentée.

1.2 OBJECTIFS

Les objectifs de ce travail sont :

1. Analyser l'évolution des indicateurs de qualité du programme valaisan
2. Identifier les déterminants de la qualité des lectures radiologiques
3. Quantifier les cancers diagnostiqués entre deux participations (cancers d'intervalle), notamment en termes de qualité des lectures et de performances du programme

^a Europe Against Cancer. *European guidelines for quality assurance in breast cancer screening and diagnosis*. 4th edition, Perry N *et al.* eds, 2006, Luxembourg: Office for Official Publications of the European Communities, 416 p.

^b Ordonnance Fédérale du 23 juin 1999 sur la garantie de la qualité des programmes de dépistage du cancer du sein réalisés par mammographie.

^c Bulliard J-L, Levi F. Dépistage du cancer du sein dans les cantons du Jura et de Neuchâtel, ainsi que dans le Jura bernois: évaluation du programme intercantonal 2005-2010. Lausanne: IUMSP, 2011 (Raisons de santé, no 180).

Bulliard J-L, Levi F. Evaluation épidémiologique du programme valaisan de dépistage du cancer du sein: 1999-2009. Lausanne: IUMSP, 2010 (Raisons de santé, no 167).

Bulliard J-L, Levi F. Evaluation épidémiologique du programme vaudois de dépistage du cancer du sein, 1999-2006. Lausanne: IUMSP, 2007 (Raisons de santé, no 136).

^d Bulliard J-L, Zwahlen M, Fracheboud J. Dépistage par mammographie en Suisse, Année 2010 / Mammografiescreening Schweiz, 2010 (rapport disponible: www.depistage-sein.ch).

^e Bulliard J-L, Levi F. Performance des lectures radiologiques dans le programme valaisan de dépistage du cancer du sein, Lausanne: IUMSP, 2006.

4. Etablir le profil pronostique des cancers d'intervalle et déterminer les facteurs caractérisant ces cancers par rapport aux cancers dépistés dans le cadre du programme.

Les résultats sont, autant que possible, mis en parallèle avec les mesures prises par la direction du Programme valaisan afin d'en estimer les effets éventuels.

1.3 STRUCTURE ET CONTENU

Ce rapport se compose de 6 chapitres et de 3 annexes.

- Le *chapitre 1* précise le **cadre** et les **objectifs** de cette évaluation thématique
- Le *chapitre 2* détaille **l'évolution de l'activité du Programme** entre 1999 et 2012 et les **mesures prises pour améliorer la qualité des lectures**
- Le *chapitre 3* traite de **l'évolution des performances** du Programme, en particulier les indicateurs de qualité (Objectif 1)
- Le *chapitre 4* examine les **déterminants de la qualité des lectures** (Objectif 2)
- Le *chapitre 5* traite des **cancers d'intervalle** du Programme (Objectifs 3 et 4)
- Le *chapitre 6* présente les **conclusions** et formule quelques **recommandations**
- La première annexe fournit un **glossaire** des termes principaux utilisés dans ce rapport
- La deuxième annexe décrit sous forme d'un tableau synoptique les principales mesures d'amélioration prises par la direction du Programme ces 10 dernières années
- La dernière annexe détaille les résultats des analyses tendanciennes de performance.

2 EVOLUTION DE L'ACTIVITE DU PROGRAMME, 1999-2012

2.1 PARTICIPANTES ELIGIBLES

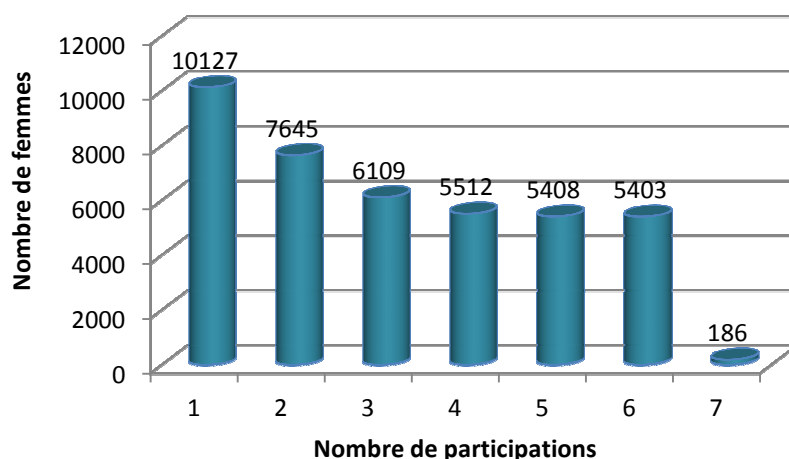
Les critères d'éligibilité démographiques et médicaux en vigueur au niveau national ont été appliqués aux participantes recensées par le Centre valaisan de dépistage (CVD)^f. Ces critères, vérifiés à chaque tour de dépistage, correspondent largement à ceux utilisés lors des évaluations antérieures. Ils permettent ainsi des comparaisons temporelles et géographiques des résultats de manière homogène.

Afin d'être éligible pour cette évaluation, une femme doit au moment de son invitation:

- Etre âgée de 50 à 69 ans^g
- Etre vivante et domiciliée (adresse connue) en Valais
- Etre sans antécédent de cancer du sein
- Ne pas porter de prothèse mammaire
- Ne pas être gravement atteinte dans sa santé

Ce rapport porte sur 126'552 mammographies réalisées auprès de 40'390 femmes éligibles entre octobre 1999 et novembre 2012. Une majorité de participantes (75%) a effectué plusieurs (entre 2 et 7) mammographies (Figure 1).

Figure 1 Nombre de participations par femme au Programme valaisan de dépistage (octobre 1999 à novembre 2012)



^f Bulliard J-L, Zwahlen M, Fracheboud J. Dépistage par mammographie en Suisse, Année 2010 / Mammografiescreening Schweiz, 2010 (rapport disponible: www.depistage-sein.ch).

^g Les femmes dans leur 50ème année sont incluses.

2.2 REALISATIONS ET LECTURES DES MAMMOGRAPHIES, 1999-2012

Le nombre de mammographies réalisées a cru régulièrement entre 1999 et 2012, de l'ordre de 4,9% par an. Il a franchi le cap des 10'000 mammographies en 2007, pour une moyenne annuelle de 9629 (Tableau 1 et Figure 2).

Tableau 1 Statistiques annuelles d'activité du Programme valaisan de dépistage, 1999 - 2012

| Moyennes annuelles | Nombre |
|--|-----------------|
| Mammographies réalisées | 9629 |
| Centres / Instituts de radiologie accrédités | 11 |
| Radiologues accrédités | 23 |
| dont: 2 ^e lecteurs | 10 |
| 3 ^e lecteurs | 5 |
| Lectures par radiologue* (IIQ) | 863 (338-1129) |
| Lectures par 1 ^e lecteur* (IIQ) | 447 (276-551) |
| Lectures par 2 ^e lecteur* (IIQ) | 1266 (960-1469) |

IIQ: Intervalle interquartile, soit la dispersion des 50% du nombre de lectures « centrales » obtenue après exclusion du plus petit et du plus grand quart du nombre de lectures par radiologue.

* Les lectures réalisées hors programme et chez des femmes inéligibles (cf. section 2.1) ne sont pas considérées.

Le nombre de radiologues 1^e lecteurs (R1) accrédités par le Programme a augmenté jusqu'en 2008, où il a atteint son maximum (29), avant de progressivement diminuer pour une moyenne annuelle de 23 radiologues actifs. Après une phase de démarrage (1999-2001), le nombre de 2^e lecteurs (R2) s'est stabilisé entre 2001 et 2005 (11 à 12 radiologues) avant de baisser graduellement (Figure 2).

Les R1 ont en moyenne interprété 447 mammographies par an tandis que les R2 en ont interprété 1266, ce qui représente un volume individuel de 863 lectures par an (Tableau 1). A noter que la grande disparité entre les volumes de lecture des R1 se réduit puisque 4 des 5 radiologues qui ont interprété le moins de mammographies (<250/an) n'étaient plus actifs en 2012. Aujourd'hui, la moitié des R1 interprètent plus de 500 mammographies par an.

Tous les R2 ont interprété plus de 1000 mammographies par an, à l'exception de 4 radiologues qui ne sont aujourd'hui plus actifs. Contrairement aux 1^e lectures, la charge des 2^e lectures est bien répartie avec 7 des 8 R2 qui ont lu en 2012 entre 1180 et 1477 mammographies. Le radiologue le plus actif du programme valaisan interprète en moyenne 3242 mammographies par année.

Afin de mieux apprécier l'effet du nombre de lecteurs accrédités au fil du temps, qui résulte de décisions organisationnelles, le nombre annuel de mammographies interprétées par type de lecteur a été examiné (Figure 3). Le volume de lectures des R2 a doublé en l'espace de 10 ans. Il dépasse 1000 depuis 2007 et atteint environ 1500 en 2012. Cet accroissement quasi linéaire correspond à 74 mammographies supplémentaires chaque année. Le volume individuel des R1 n'a que faiblement augmenté (10 mammographies de plus par an) et atteint les 500 lectures par an en 2012. Un accroissement continu du volume de lectures est observé depuis 2008.

Figure 2 Evolution du nombre de mammographies et de lecteurs (premiers et deuxièmes) du Programme valaisan de dépistage (10/1999 - 09/2012)

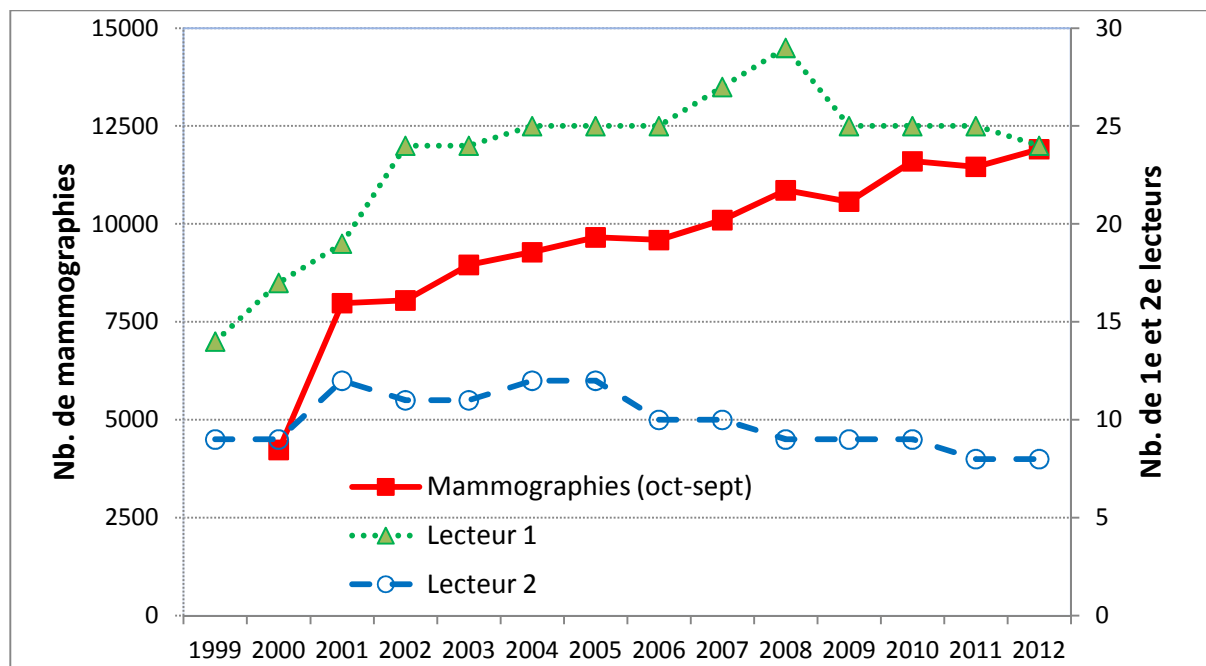
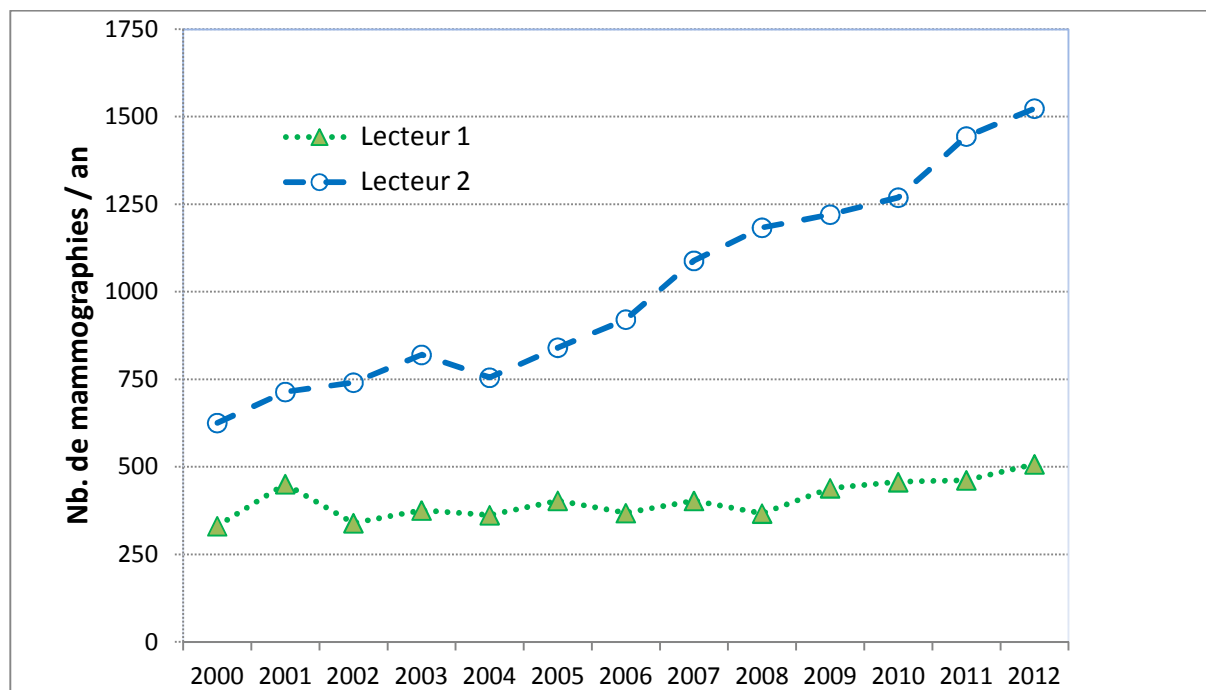


Figure 3 Evolution du nombre annuel de mammographies par type de lecteurs dans le Programme valaisan de dépistage, 2000 – 2012.



2.3 ACTIONS PRISES PAR LE PROGRAMME

Une synthèse des diverses mesures et actions prises par la direction du Programme valaisan en vue d'améliorer la qualité des lectures est fournie au Tableau 9 (Annexe 2). Ces mesures ont débuté en 2003 suite à une évaluation portant sur la période 1999-2002 et indiquant un taux de reconvoction légèrement au-delà des normes recommandées en tour prévalent. Cette observation, confirmée les années suivantes^h, a conduit à des interventions individuelles et à des mesures de sensibilisation de l'ensemble des radiologues en 2006, avec un suivi détaillé réalisé par la direction du Programmeⁱ. **Les interventions individuelles de 2006 ainsi que la mesure organisationnelle visant à contenir dès 2007-8 les accréditations de radiologues, surtout en tant que R2, sont les actions les plus significatives entreprises.** Contrairement aux mesures individuelles, ciblées et bien définies dans le temps, celle visant les nouvelles accréditations aura un impact graduel sur le volume annuel de lectures, rendant plus délicate l'identification de tout effet sur la qualité du programme.

L'impact des autres actions de sensibilisation apparaît, a priori, difficile à estimer. En effet, ces actions ont souvent été ponctuelles (annuelles), n'ont pas touché tous les radiologues (p.ex. participation aux colloques de formation continue, rôle de la thématique abordée, etc.), peuvent avoir influencé différemment les radiologues (p.ex. statistiques de performance individuelle comparée à la moyenne des radiologues) et n'intègrent pas d'autres éléments susceptibles d'influencer l'interprétation des lectures (p.ex. changement d'équipement, autres cours de formation continue, etc.). En outre, les mesures de sensibilisation générale à la qualité des lectures et les interventions organisationnelles ne ciblent qu'indirectement certains indicateurs de performance (taux de reconvoction, taux de détection et taux de cancer d'intervalle), par exemple en associant une augmentation du volume individuel de lectures à un gain qualitatif dans l'interprétation des lectures.

2.4 COMMENTAIRES ET CHOIX ANALYTIQUES

Evolution de l'activité

L'évolution de l'activité du programme (Figures 2 et 3) montre clairement 2 phases: une première, de démarrage (1999-2001), correspondant au 1^{er} tour d'invitation du programme. Cette phase initiale est caractérisée par une forte croissance des mammographies réalisées et du nombre de radiologues accrédités.

Dès 2002, le volume d'activité croît de manière plus régulière, avec une stabilisation du nombre des lecteurs, surtout des R2. Cette seconde période est plus homogène, tant par la structure par âge des participantes que sur le plan organisationnel. Toutes les mesures et actions prises par le programme l'ont été lors de cette phase.

Afin de cerner d'éventuelles variations annuelles dans l'évolution des performances susceptibles d'être associées aux actions menées par la direction du Programme, les analyses se sont concentrées sur la période 2002-12. L'inclusion des effets induits par la forte croissance initiale du Programme (1999-2001) pourrait occulter les variations temporelles intrinsèques à la période d'intérêt pour cette étude.

Afin d'identifier de façon objective toute fluctuation « anormale » dans les tendances annuelles, un outil statistique dédié a été utilisé (méthode Joinpoint). L'application de cette méthode a permis de révéler un accroissement significativement plus important du volume de lecture des R1 dès 2008 par rapport à la

^h Bulliard J-L, Levi F. Performance des lectures radiologiques dans le programme valaisan de dépistage du cancer du sein, Lausanne: IUMSP, 2006.

ⁱ Programme valaisan de dépistage du cancer du sein. Evolution statistique des lectures radiologiques 2005 – 2007 (rapport interne), 2007.

^j La méthode statistique dite « JoinPoint » identifie des changements significatifs dans une tendance linéaire et décompose cette tendance en segments linéaires. Basée sur la théorie des permutations, la méthode JoinPoint fournit une quantification de la tendance (changement % annuel) pour chaque segment linéaire identifié (cf. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med* 2000;19:335-51 (correction: 2001;20:655)).

période 2002-7 (augmentation de 25 lectures individuelles par an entre 2008 et 2012 contre 6 entre 2002 et 2007). Par ailleurs, aucun changement significatif dans la croissance du volume individuel de lectures des R2 n'a été détecté entre 2002 et 2012. Ces élaborations statistiques corroborent les tendances illustrées dans la Figure 3.

Volume de lectures

Dans un contexte où le nombre de mammographies augmente au fil du temps, un « contrôle » prospectif plus strict, par la direction du Programme, du nombre de radiologues accrédités ou acceptés comme nouveaux R2 a permis de graduellement concentrer les lectures sur moins de radiologues. Cet effet s'est traduit par une augmentation constante du volume individuel, surtout manifeste pour les R2 puisque la charge des lectures y est distribuée entre moins de radiologues que pour les 1^e lectures (Figure 3). Si cette tendance se poursuit, en 4 ans le nombre moyen de mammographies interprétées par les R2 aura cru d'environ 300. Cette progression est notable au vu des mesures prises.

La taille du programme n'a pas permis d'assurer 500 lectures annuelles à tous les R1 puisque le ratio entre les nombres de radiologues accrédités et de mammographies réalisées est globalement inférieur à 1:500 (Figure 2). Un seuil de 500 lectures annuelles en moyenne serait réalisable en 2012 avec une distribution uniforme des 1^e lectures, une condition toutefois difficile à satisfaire en pratique vu que la 1^e lecture est effectuée au centre de radiologie où s'est rendue la femme. Cette évolution croissante du volume des R1 peut être largement attribuée à l'application d'une politique plus stricte d'accréditation de nouveaux lecteurs.

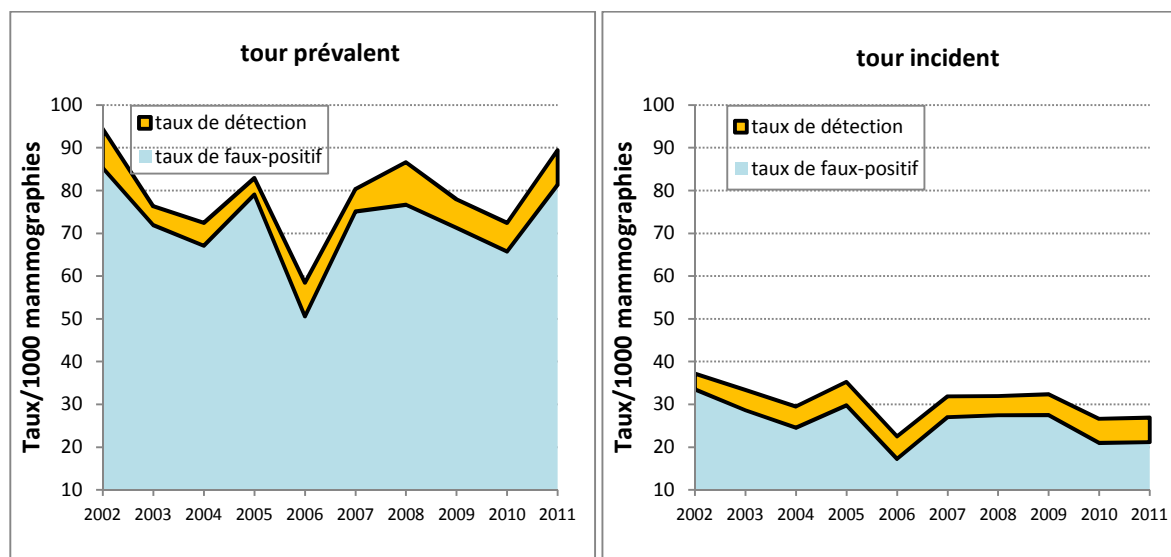
3 EVOLUTION DES PERFORMANCES DU PROGRAMME

Au total 101'039 mammographies ont été réalisées entre 2002 et 2011, dont 3 sur 4 (76,5%) en tour incident. Ainsi, les indicateurs en tour incident portent non seulement sur un collectif plus homogène (des femmes pour lesquelles les radiologues disposent de clichés comparatifs dans le cadre du Programme) mais ils y sont estimés de manière plus précise.

3.1 EVOLUTION DU TAUX DE RECONVOCATION

Le taux de reconvoication indique la proportion des femmes dépistées qui sont rappelées pour des examens complémentaires suite à l'identification sur leurs mammographies d'une anomalie jugée suspecte (voir glossaire, Annexe 1). Ces investigations diagnostiques confirmeront ou infirmeront la suspicion radiologique de cancer. Les taux annuels de reconvoication du programme valaisan ventilés selon le résultat final des investigations (détection d'un cancer ou examen faux positif) sont présentés dans la Figure 4 séparément pour les tours prévalent (1^{ère} participation) et incident (participations subséquentes). L'analyse a porté sur la décennie 2002-11 puisque la moitié environ des résultats des examens complémentaires de 2012 n'était pas encore disponible (58,7% et 42,4% en tour prévalent et incident, respectivement).

Figure 4 Evolution entre 2002 et 2011 du taux de détection et de faux positif du Programme valaisan de dépistage par mammographie en tour prévalent et incident



Pour l'ensemble de la période 2002-11, le taux de reconvoication est de 80,4/1000 mammographies en tour prévalent (normes européennes: < 50-70/1000) et de 30,5/1000 en tour incident (normes européennes: < 30-50/1000) (Tableau 2). Même corrigé pour l'âge^k (taux de reconvoication ajusté: 76,8/1000), ce taux reste supérieur aux recommandations européennes en première participation. Les taux de détection de cancer en tours prévalent et incident sont respectivement de 6,9/1000 et 4,8/1000, de sorte que les faux positifs représentent, respectivement, 73,5 et 25,4 résultats pour 1000 examens. La valeur prédictive positive (VPP) des mammographies de dépistage réalisées entre 2002 et 2011 est ainsi de

^k La distribution de la densité mammaire des participantes diffère entre le tour prévalent et les tours incidents, surtout à cause de l'âge. Une pondération statistique sur l'âge permet de comparer des groupes de distributions de densité (et d'âge) différentes.

Tableau 2 Valeurs (globales et récentes) et évolution des principaux indicateurs de qualité du Programme valaisan de dépistage du cancer du sein

| Indicateur | Tour | Période | | Tendance annuelle* | Commentaires |
|--|-----------|--------------------|--------------------|--------------------|--|
| | | globale 2002-11 | récente 2009-11 | | |
| Taux de reconvoction (/1000 mammographies) | Prévalent | 80,4 | 79,7 | -0,8% | Tendance non significative** |
| | Incident | 30,5 | 28,7 | -2,2% | Tendance non significative** |
| Taux de détection (/1000 mammographies) | Prévalent | 6,9 | 7,1 | +0,4% | Tendance non significative** |
| | Incident | 5,1 | 5,4 | - | Tendance fluctuante (2002-5:↗, 2006-8:↘, 2009-11:↗) |
| Cancers in situ (/100 cancers dépistés) | Prévalent | 21,3 | 31,1 | +5,0% | Tendance non significative** |
| | Incident | 13,3 | 13,3 | -0,7% | Tendance non significative** |
| Cancers invasifs ≤1cm (/100 cancers invasifs) | Prévalent | 32,5 | 30,0 | +7,8% | Tendance non significative** |
| | Incident | 22,4 | 27,4 | +5,2% | Tendance non significative** |
| Cancers de stade avancé*** (/100 cancers dépistés) | Prévalent | 32,7 | 35,7 | - | Tendance fluctuante (2002-5:↗, 2005-7:↘, 2007-9: ↗, 2009-11:↘) |
| | Incident | 33,9 | 29,4 | -6,2% | Tendance significativement à la baisse |

* Résultats obtenus par application de la méthode JoinPoint (cf. section 2.4)

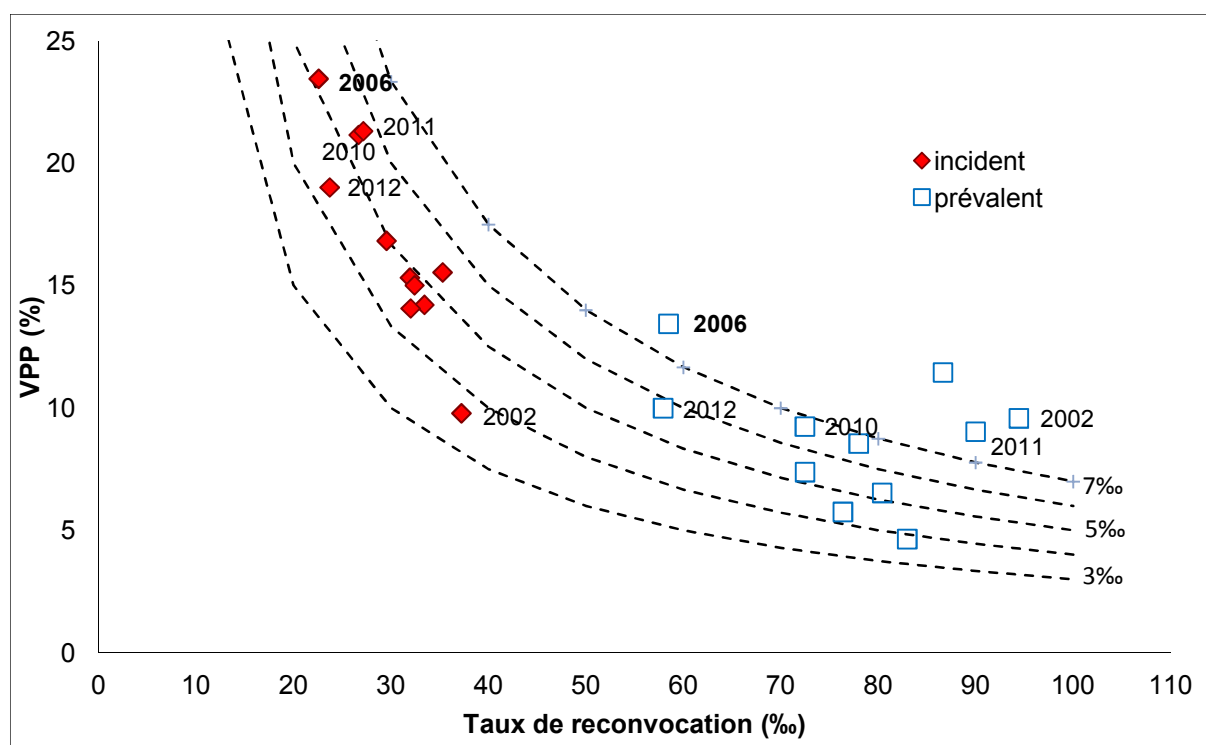
** Ceci indique que la tendance annuelle peut être statistiquement considérée comme nulle (stabilité)

*** Les cancers de stade inconnu sont exclus (soit 17 cancers prévalents et 38 cancers incidents)

8,9% en tour prévalent (1 cancer pour 11 mammographies suspectes) et 16,8% en tour incident (1 cancer pour 6 mammographies suspectes). A noter que la VPP est plus élevée en 2006 (13,4% et 23,5% en tour prévalent et incident, respectivement), année où une baisse importante du taux de reconvoation mais pas du taux de détection est observée (Figure 4).

Pris isolément, le taux de reconvoation est un indicateur d'intérêt limité puisque des taux de reconvoation différents peuvent conduire à un même taux de détection (voir Figure 5: les courbes en pointillés indiquent la relation entre le taux de reconvoation et un taux de détection donné (de 3‰ à 7‰): plus les performances se rapprochent du coin supérieur gauche du graphique, meilleures elles sont puisque la VPP et le taux de détection augmentent pour un taux de reconvoation donné). La Figure 5 illustre et confirme 1) une plus grande disparité des résultats et de moins bonnes performances en tour prévalent qu'incident, 2) une tendance vers de meilleures performances pour les années récentes (les points pour 2010, 2011 et 2012) et 3) de meilleures performances en absolu pour l'année 2006.

Figure 5 Taux de reconvoation et VPP par année et type de tour de dépistage, 2002-12



3.2 ANALYSE DES TENDANCES DES INDICATEURS DE QUALITE

Le taux de reconvoation montre une évolution linéaire avec une diminution modeste (non statistiquement significative) de 0,8%/an en tour prévalent et de 2,2%/an en tour incident (Tableau 2). Cette baisse annuelle de 2,2% reflète une diminution de 25% environ sur 10 ans du taux de reconvoation (de 4% en 2002 à 3% en 2011). Comme aucune diminution concomitante du taux de détection n'a été enregistrée, cette baisse s'est essentiellement répercutée sur les résultats faux positifs. A noter la volatilité des tendances du taux de détection en tour incident, où une baisse ponctuelle a été identifiée entre 2006 et 2008 (Tableau 2 et Figure 9, Annexe 3).

Le taux - globalement stable - de détection de cancer ne signifie pas forcément que cette évolution soit la même pour les cancers dépistés précocement et tardivement. Un examen approfondi révèle une tendance dans le Programme à détecter une proportion de moins en moins importante de cancers de stade avancé, surtout en tour incident. Cette baisse, statistiquement significative, de 6,2%/an explique la moindre

proportion de cancers de stade avancé depuis 2009 par rapport à l'ensemble de la période 2002-11 (Tableau 2). En revanche, on observe une proportion croissante (non significative) de cancers dépistés à pronostic favorable (taille ≤ 1 cm ou cancers in situ), particulièrement en première participation. Il est intéressant par ailleurs de relever que la taille médiane des cancers dépistés est identique en tour prévalent et incident (14 mm) et qu'elle fluctue peu au fil des ans.

3.3 COMMENTAIRES

La tendance générale est confortante avec une amélioration globale de la qualité des lectures. En tour incident, où les indicateurs peuvent être estimés plus fiablement, la proportion de cancers de stade avancé a baissé significativement au fil des ans et se rapproche des normes européennes (29,4% pour la période 2009-11 contre 25% au maximum recommandé). Cette évolution favorable se manifeste aussi par une légère baisse du taux de reconvoation (et du taux de faux positifs) et une modeste hausse dans la proportion de cancers dépistés de taille inférieure ou égale à 1 cm. La proportion de cancers de petite taille approche d'ailleurs le référentiel européen (27,4% pour la période 2009-11 contre 30% au minimum recommandé). La plupart des tendances restent cependant modestes et ne peuvent exclure comme explication des fluctuations aléatoires (tendance non significative). La fraction élevée et croissante des cancers in situ en première participation mériterait une surveillance épidémiologique puisque cet indicateur est un indice potentiel de sur-détection.

Il est difficile d'attribuer cette évolution aux mesures prises par le Programme. Seule la forte baisse du taux de reconvoation en 2006 et l'augmentation concomitante de la VPP de la mammographie coïncident largement avec les interventions individuelles réalisées par la direction du programme. Ces effets bénéfiques sont ponctuels et suggèrent que des actions répétées sont nécessaires pour impacter plus durablement sur la qualité radiologique d'un programme. Les actions individualisées, tel un entretien avec chaque radiologue, apparaissent les plus efficaces et prometteuses.

4 DETERMINANTS DE LA QUALITE DES LECTURES

4.1 CONSIDERATIONS PRELIMINAIRES

L'analyse des déterminants de la qualité des lectures porte sur les mammographies réalisées entre 2002 et 2012 afin de s'affranchir de la phase de démarrage du programme pendant laquelle le volume de lectures était comparativement faible et peu représentatif des conditions d'activité actuelle du Programme (voir section 2.4).

Les résultats des lectures d'un radiologue sont essentiellement influencés par 3 types de facteurs:

- 1) les caractéristiques des patientes (niveau individuel de risque, densité du tissu mammaire, etc.),
- 2) la qualité technique des clichés (appareil et technologie utilisés, qualité des clichés des TRM)
- 3) ses performances intrinsèques.

Une exploration du « case-mix » des mammographies interprétées par chaque radiologue révèle de substantielles différences dans le degré de difficulté de lecture (p.ex. de 13 à 48% de seins denses (BIRADS 3 ou 4), une différence significative dans l'âge des patientes, etc.). Cette situation peut se produire lorsque, par exemple, les cas délicats ou les patientes à risque sont référés préférentiellement vers un ou quelques centres de radiologie particuliers. Ces différences influent sur les performances individuelles de lecture et faussent toute comparaison directe entre radiologues. Le recours à un modèle mathématique complexe permettrait de prendre en compte, du moins partiellement, ces facteurs et leurs interactions. Cependant, la traduction opérationnelle des résultats s'avérerait difficile sur le plan pratique. A des fins de comparaisons individuelles, qui n'est pas l'objectif de cette analyse, une solution reste la relecture d'un nombre limité de mammographies, dans les mêmes conditions, par tous les radiologues. Cette approche nécessiterait la mise sur pied d'une étude *ad hoc*, idéalement dans un cadre inter-cantonal. Afin d'évaluer les déterminants des interprétations radiologiques à un niveau agrégé, les facteurs potentiellement influant la qualité des lectures ont été examinés (voir section 4.3 ci-dessous).

4.2 RESULTATS DES LECTURES MULTIPLES

Les Figures 6 et 7 détaillent pour les tours prévalent et incident les résultats de la lecture multiple des 111'983 mammographies réalisées entre 2002 et 2012. On observe globalement que:

- Les conclusions de la double lecture concordent plus de 9 fois sur 10 (92,5%)
- La L2 confirme 22 fois sur 100 la positivité d'une L1 et 97 fois sur 100 sa négativité
- Les positifs en L1 et L2 divergent 1 fois sur 2 (51,0%) sur la lésion identifiée, nécessitant une L3 (8,5% de toutes les L3)
- Une L3 est requise dans moins de 1 cas sur 10 (8,2%); le recours à une L3 est cependant 2 fois plus fréquent en tour prévalent qu'incident (13,9% contre 6,5%)
- Sur 100 divergences de la double lecture (L1+L2- ou L1-L2+), la L3 confirme 45 fois une L2 positive et 33 fois une L1 positive

Lorsque les renseignements transmis par un des 2 lecteurs étaient incomplets, généralement sans conclusion, le CVD a demandé une 3^e lecture. Ces rares cas (n=55) apparaissent comme des L3 après une double lecture négative dans les Figures 6 et 7 (en traitillés).

Figure 6 Résultats des lectures multiples du Programme valaisan en tour prévalent, 2002-2012

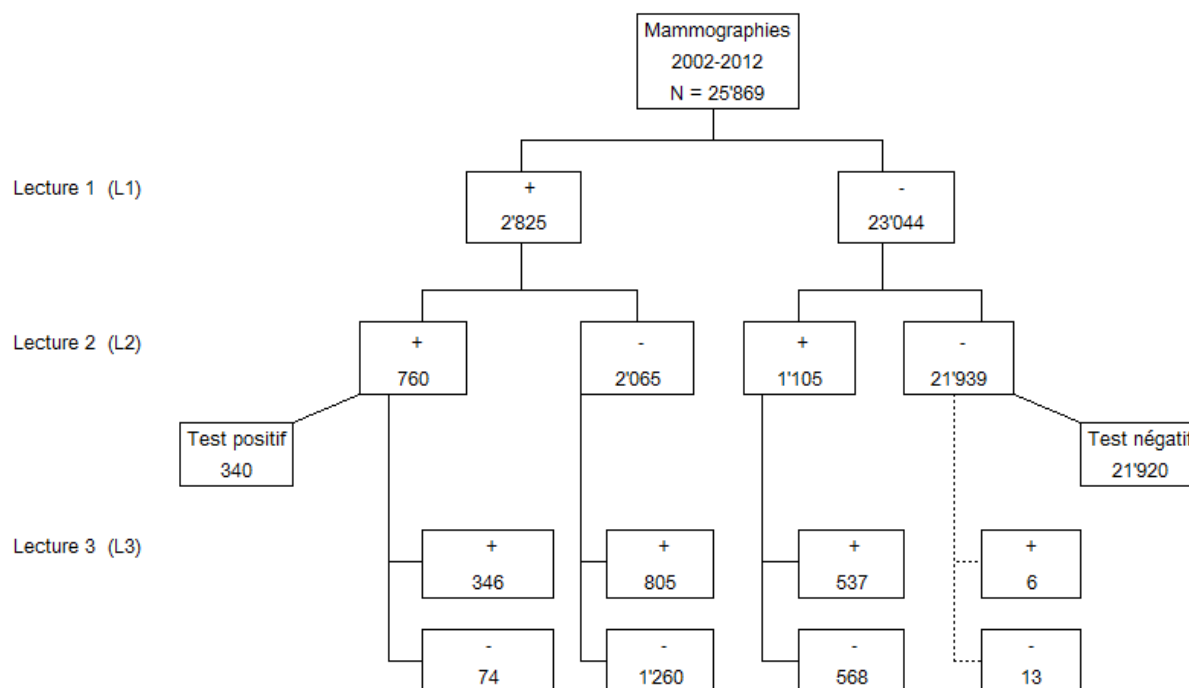
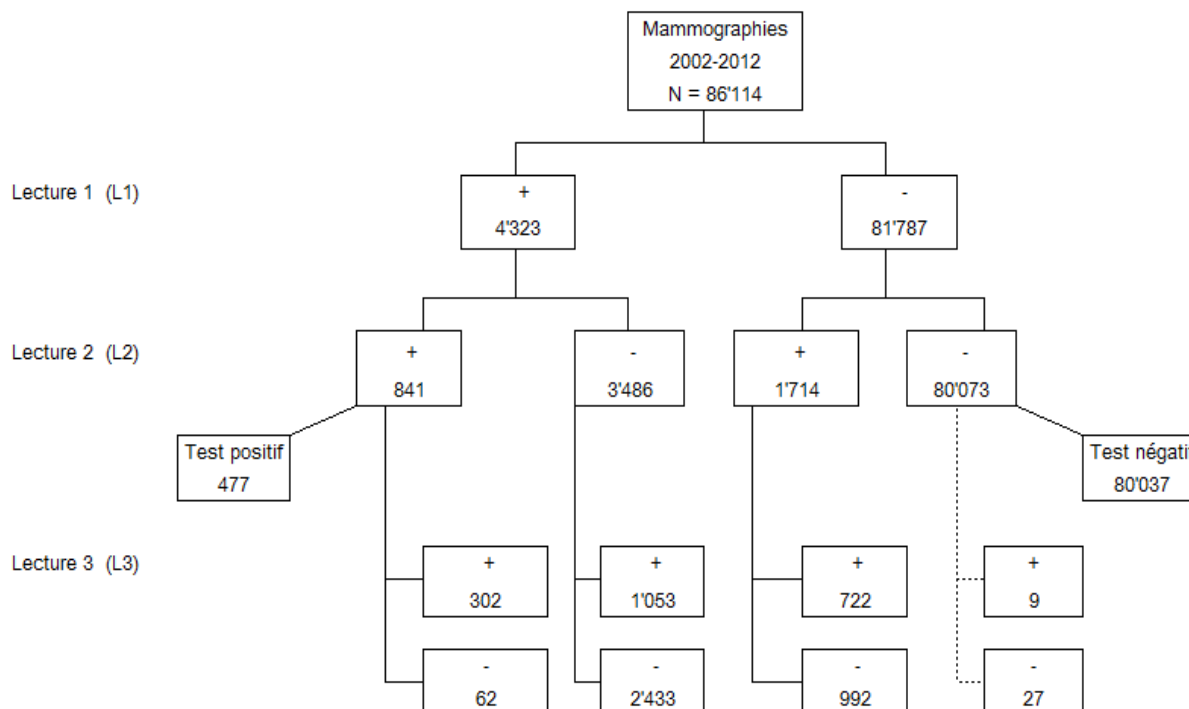


Figure 7 Résultats des lectures multiples du Programme valaisan en tour incident, 2002-2012



4.3 DETERMINANTS DES INTERPRETATIONS RADIOLOGIQUES

Six facteurs susceptibles d'influer la qualité de l'interprétation des mammographies ont été examinés:

- 1) Type de tour (prévalent ou incident)
- 2) L'année de réalisation de la mammographie
- 3) Contexte de lecture (première ou deuxième)
- 4) Type de lecteur (premier ou second)
- 5) Volume de lectures individuel (annuel ou cumulé)
- 6) L'équipement radiologique (technologie analogique ou numérique)

L'apport de clichés comparatifs (tour incident) sur la qualité des interprétations radiologiques a pu largement être confirmé. Ainsi, pour les 5 autres facteurs considérés, les performances de lecture étaient systématiquement meilleures en tour incident que prévalent: baisse du taux de faux positif sans impact sur le taux de cancer non vus et hausse de 35 à 70% de la VPP. Le gain de qualité lié à la présence de clichés antérieurs était le plus manifeste pour les radiologues 2^e lecteurs (VPP de 7,7% et 13,1% en L2 en tours prévalent et incident, respectivement). L'effet du type de tour semblant agir indépendamment des autres facteurs, les analyses ci-dessous ont combiné tous les tours de dépistage.

Le passage à la technologie numérique dans les centres de radiologie s'est étalé entre mars 2006 et juin 2012. Comme les L2 au CVD se faisaient jusqu'en avril 2012 sur console analogique, l'effet du passage au numérique sur la qualité des interprétations radiologiques n'a pu être mesuré que pour les L1. Cette transition technologique ne semble pas avoir eu d'effet majeur sur les performances en L1: les taux de mammographies positivées (analogique: 63,8‰ ; numérique: 61,9‰), de détection (4,5‰ dans les 2 cas) et de cancers dépistés qui ont échappé au 1^e lecteur (analogique: 0,8‰ ; numérique: 0,6‰) sont comparables. Ainsi, les VPP sont similaires pour les L1 analogiques (7,1%) et numériques (7,2%).

Les taux de détection, de faux positif, de cancers non vus ainsi que la VPP selon le contexte de lecture, le type de lecteur et le volume individuel de lectures sont détaillés au Tableau 3. De manière générale, les performances sont légèrement meilleures en 2^e qu'en 1^e lecture: on y positive moins souvent (38,9‰ vs 62,9‰), surtout à tort (taux de faux positif: 34,7‰ vs 58,4‰), avec, en contrepartie, un déficit mineur de cancers non dépistés (0,3‰). Ainsi, la VPP en L2 est de 10,8% contre 7,2% en L1.

Comparer les résultats des 2^e lecteurs lorsqu'ils effectuent des L1 ou des L2 permet de déterminer si cette différence de performance est plutôt intrinsèque aux radiologues ou liée au contexte de lecture. On constate que les performances des 2^e lecteurs en L1 sont très proches de celles des 1^e lecteurs (VPP de 7,3% contre 7,0% pour les 1^e lecteurs) et inférieures à leurs performances en L2 (VPP de 10,8%). Le contexte de lecture apparaît donc déterminant. Ce constat est corroboré par les résultats selon le volume de mammographies interprétées: à volume de lectures comparables (annuel ou cumulé), les performances sont souvent supérieures en L2 qu'en L1 (voir colonnes VPP du Tableau 3).

Les radiologues ayant interprété le plus de mammographies dans le cadre du programme ont généralement un taux de faux positif plus bas - donc une meilleure spécificité - et ratent moins de cancers vus par leurs pairs. Toutefois, ils tendent à détecter moins de cancers, surtout en L1, de sorte que la VPP n'augmente pas avec le volume de lectures. La seule exception est l'observation en L2 d'une VPP qui croît avec le volume cumulé de lectures entre 2002 et 2012. A noter qu'aucun 2^e lecteur n'a interprété moins de 500 mammographies annuelles et moins de 2000 mammographies au cours de ces dix dernières années.

La qualité des lectures, résumée par la VPP, montre une légère amélioration (non significative) au fil des ans (voir Figure 10, Annexe 3). Cette évolution est plus prononcée en L2 qu'en L1 (0,4% contre 0,1% de croissance annuelle de la VPP). A noter que les VPP les plus élevées, tant en L1 qu'en L2, ont été observées en 2006.

Tableau 3 Taux de faux positif, de détection, de cancers non vus, et valeur prédictive positive (VPP) selon le contexte de lecture, le type de lecteur et le volume individuel de lecture, tous tours confondus (2002-12)

| Facteurs | Faux positif (‰) | | Détection* (‰) | | Ca. non vus (‰) | | VPP* (%) | |
|-----------------------------------|------------------|-------------|----------------|------------|-----------------|------------|------------|-------------|
| | L1 | L2 | L1 | L2 | L1 | L2 | L1 | L2 |
| | n=6504 | n=3867 | n=502 | n=470 | n=79 | n=111 | | |
| Type de lecteur | | | | | | | | |
| 1 ^{er} lecteur | 62,6 | - | 4,7 | - | 0,7 | - | 7,0 | - |
| 2 ^e lecteur | 53,6 | 34,7 | 4,2 | 4,2 | 0,7 | 1,0 | 7,3 | 10,8 |
| Volume annuel de lectures* | | | | | | | | |
| moins de 500 | 53,4 | - | 4,6 | - | 1,0 | - | 7,9 | - |
| 500 – 999 | 72,9 | 45,7 | 5,1 | 3,9 | 0,6 | 1,1 | 6,6 | 7,8 |
| 1000 – 1999 | 46,8 | 30,1 | 3,9 | 4,0 | 0,7 | 1,1 | 7,6 | 11,8 |
| 2000 ou plus | 45,5 | 43,8 | 3,1 | 5,0 | 0,3 | 0,6 | 6,5 | 10,3 |
| Volume cumulé de lectures* | | | | | | | | |
| moins de 2000 | 72,4 | - | 4,5 | - | 0,8 | - | 5,8 | - |
| 2000 – 4999 | 55,1 | 51,6 | 4,8 | 4,2 | 0,9 | 1,4 | 8,0 | 7,4 |
| 5000 – 9999 | 66,1 | 29,9 | 5,1 | 3,6 | 0,6 | 0,8 | 7,1 | 10,6 |
| 10'000 ou plus | 53,1 | 34,5 | 4,0 | 4,4 | 0,6 | 1,0 | 7,0 | 11,2 |
| Total | 58,4 | 34,7 | 4,5 | 4,2 | 0,7 | 1,0 | 7,2 | 10,8 |

L1 : 1^e lecture, L2 : 2^e lecture

* On considère que la lésion positivée par le radiologue a permis d'initier les investigations qui ont mené au diagnostic de cancer (même si la correspondance exacte entre la lésion et le cancer n'est pas avérée). Il en découle une probable légère surestimation du taux de détection et de la VPP.

4.4 COMMENTAIRES

Cette analyse de la qualité des lectures et de ses déterminants est la plus aboutie à ce jour pour un programme suisse. Elle confirme l'importance du type de tour et du contexte de lecture, mis en évidence lors d'une première évaluation de la qualité des lectures du programme valaisan en 2006¹. Les résultats attestent aussi que les mammographies interprétées par les radiologues sont réparties avec des degrés de difficulté hétérogènes, que la qualité des lectures s'améliore graduellement, est indépendante de la technologie (analogique ou numérique) et n'est pas fortement associée au volume individuel de lecture. Par ailleurs, le taux de faux positif capture la majeure partie des différences globales de performance entre radiologues, notamment de VPP. Ce dernier élément, avec la prépondérance très nette du contexte de lecture sur le type de lecteurs (1^e ou 2^e) sont les déterminants de la qualité de l'interprétation iconographique sur lesquels le programme pourrait directement influencer.

Plusieurs facteurs sont susceptibles d'expliquer les différences de performance entre la L1 et la L2:

¹ Bulliard J-L, Levi F. Performance des lectures radiologiques dans le programme valaisan de dépistage du cancer du sein, Lausanne: IUMSP, 2006.

1. Le contexte des L2, réalisées au CVD, s'avère vraisemblablement plus favorable que celui des L1, effectuées au cabinet, où le risque d'interruption pour des sollicitations diverses est plus fréquent.
2. En L1, le radiologue est plus facilement porté à se placer dans une approche individuelle « médecin-patient », qui favorise la sensibilité, plutôt que dans une démarche purement de santé publique, typique du contexte de L2, et qui favorise la spécificité. Les résultats (meilleure spécificité mais moins bonne sensibilité en L2 qu'en L1) corroborent cette hypothèse.
3. La conséquence de l'interprétation radiologique peut être perçue différemment selon le contexte de lecture (sentiment de disposer d'un filet de sécurité en L1).

Contrairement au programme valaisan, la différence de performances entre 1^e et 2^e lecteurs est marquée dans les autres programmes romands. Ces programmes ont décidé dès leur démarrage de confier les L2 à un nombre restreint de radiologues, dont les performances étaient avérées, afin qu'ils puissent interpréter un volume de lectures si possible en adéquation avec les recommandations européennes (5000 par an). La faible différence de performance (à contexte égal) entre R1 et R2 du programme valaisan reflète vraisemblablement la stratégie « d'homogénéité » mise en place de ne pas opérer de distinction entre les types de lecteurs de sorte que les radiologues volontaires pour être 2^e lecteurs présentaient initialement un niveau de capacité similaire aux R1. Le léger gain de performance observé en L1 pour les R2 reflète ainsi essentiellement l'expérience supplémentaire acquise par l'interprétation d'un volume de lecture globalement supérieur. Le nombre de R2 étant encore comparativement élevé en Valais (une restriction des accréditations des R2 a débuté en 2007-8, voir Tableau 9, Annexe 2), leur volume annuel de lectures reste non seulement inférieur à celui des R2 des autres programmes romands, mais la différence de volume de lectures avec les R1 est moins marquée (rapport des volumes de lecture entre R2 et R1 d'environ 2.8, voir Tableau 1, section 2.2, alors qu'il approche 4 dans d'autres programmes). Ces résultats confirment l'effet favorable mais modeste des mesures prises au niveau du nombre de lecteurs, surtout les R2.

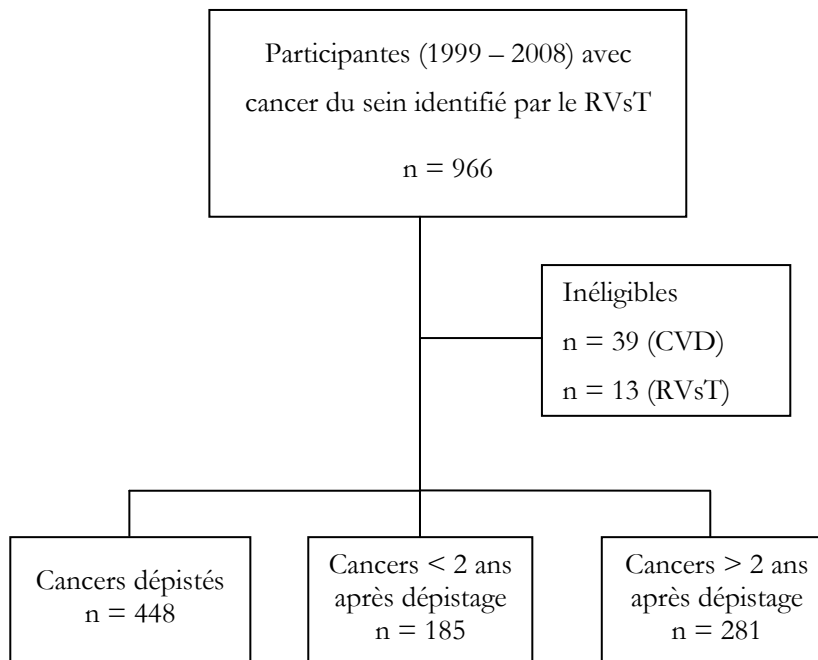
La différence peu marquée de volume de lectures entre R1 et R2 et les volumes de lecture globalement modestes pourraient partiellement expliquer pourquoi aucune association forte n'émerge entre le volume et la qualité des lectures. Les R2 ayant d'abord œuvré comme R1, les différences de volume de lectures entre radiologues sont plus prononcées pour le volume cumulé que le volume annuel. Sur une échelle cumulative, l'effet du volume de lectures sur la qualité apparaît un peu plus clairement pour les R2 avec un gain de performance au-delà de 5'000 mammographies interprétées. Un effet favorable sur la qualité des lectures est également suggéré pour un volume annuel de lectures supérieur à 1000 mammographies. Rappelons que seules les lectures effectuées dans le cadre du programme peuvent être comptabilisées et que les radiologues qui y interprètent moins de clichés lisent possiblement plus de mammographies hors programme. Un tel phénomène contribuerait à réduire les différences de qualité associée au volume de lectures.

5 CANCERS D'INTERVALLE

5.1 IDENTIFICATION

Mille et un cancers diagnostiqués entre 1999 et 2010 chez 966 participantes au Programme ont été identifiés par croisement informatique entre le Registre Valaisan des Tumeurs (RVsT) et la base de données du CVD^m. Ces cas constituent le matériel mis à disposition sous forme anonyme par le RVsT pour ce rapport. Comme l'identification et la documentation des cancers d'intervalle exigent un délai minimal de 2 ans après l'examen de dépistage, ces cancers portent sur des mammographies réalisées entre 1999 et 2008. Suite aux croisements, 39 femmes inéligibles pour l'évaluation et 13 apparemment éligibles mais ayant eu un cancer du sein avant leur dépistage ont été exclues, laissant 448 cancers dépistés et 466 cancers diagnostiqués après une mammographie de dépistage (Figure 8).

Figure 8 Flux et classification des cancers chez les participantes (1999-2008) identifiés par le RVsT



La qualité de la procédure de croisements est primordiale pour identifier de manière exhaustive et valide les potentiels cancers d'intervalle. Une évaluation de la procédure utilisée par le RVsTⁿ indique que 4,2% des cancers dépistés par le programme n'ont pas été appariés (19/450). Des changements importants dans certains dossiers (modification de la date de naissance ou du nom) ont induit des discordances qui sont à l'origine du non croisement de 16 cas pourtant identifiés lors d'évaluations antérieures (appariements de 2006 et de 2009). Ces cas ont été complétés. Ainsi, seuls 3 cancers dépistés n'ont pu être identifiés par croisement informatique (2 cas inconnus du RVsT et 1 dont le domicile pour le RVsT est hors canton).

^m Seules les participantes dûment autorisant la transmission de leurs données à des fins statistiques et d'amélioration de la qualité du programme de dépistage ont été considérées. 2010 est l'année d'incidence complète la plus récente disponible auprès du RVsT.

ⁿ Dal Maso L, Braga C, Franceschi S. Methodology used for "software for automated linkage in Italy" (SALI). *J Biomed Inform* 2001;34:387-95.

5.2 FREQUENCE

Au total, 185 cancers ont été diagnostiqués chez des participantes dans les 2 ans suivant leur dernier examen de dépistage négatif (Figure 8). La fréquence de ces cancers d'intervalle est présentée par type de tour et année après dépistage au Tableau 4.

Tableau 4 Fréquence des cancers d'intervalle (1999-2010) du Programme valaisan de dépistage par mammographie et comparaison aux normes européennes.

| | Tour | | Normes européennes | |
|---|---------------------------|---------------------------|--------------------|----------|
| | prévalent (70 cancers) | incident (115 cancers) | prévalent | incident |
| Taux de cancers d'intervalle (2 ans) (/1000 femmes dépistées) | 2,2 | 2,3 | | |
| Première année (0-11 mois) | 0,7 | 0,7 | | |
| Deuxième année (12-23 mois) | 1,5 | 1,6 | | |
| Incidence proportionnelle* (%) | | | | |
| Première année (0-11 mois) | 24,8 | 24,9 | ≤30 | ≤30 |
| Deuxième année (12-23 mois) | 50,6 | 54,7 | ≤50 | ≤50 |

* Exprimée en % du taux d'incidence « en l'absence de dépistage » (l'incidence du cancer du sein en Valais pour la période précédant le démarrage du programme (incidence 1994-98: 292,1/100'000 personnes âgées de 50 à 69 ans) a été considérée comme incidence en absence de dépistage).

La fréquence des cancers d'intervalle est similaire en tour prévalent et incident. Son incidence représente environ 40% de celle des cancers dépistés par le Programme. Le risque de cancer d'intervalle double entre la première et la deuxième année. L'incidence proportionnelle des cancers d'intervalle en première année satisfait les normes européennes. Elle est quelque peu élevée en seconde année (50,6% et 54,7% de l'incidence estimée en « absence de dépistage » en tour prévalent et incident respectivement contre un maximum recommandé de 50%). Vu le nombre modeste de cancers d'intervalle, les fréquences annuelles présentaient de trop grandes variations pour en tirer des conclusions sur des tendances éventuelles.

La connaissance des cancers d'intervalle – et par conséquent des résultats « vrais négatifs » et « faux négatifs » du dépistage – permet d'estimer la sensibilité et la spécificité du programme (Tableau 5). Tous tours confondus, la sensibilité est de 70,7% et la spécificité de 95,8%. La spécificité augmente entre les tours prévalent et incident tandis que la sensibilité diminue. La valeur prédictive positive (VPP) de la mammographie est de 11,6%: elle augmente entre les tours prévalent et incident, tandis que la valeur prédictive négative (VPN) est stable (99,8%).

Tableau 5 Sensibilité et spécificité par type de tour du Programme valaisan de dépistage par mammographie.

| Tour prévalent | | | Tour incident | | | | |
|---|---|-----|--|--------------|---|-----|-------|
| Diagnostic final de cancer | | | Diagnostic final de cancer | | | | |
| | + | - | | + | - | | |
| Mammographie | + | 204 | 2093 | Mammographie | + | 242 | 1317 |
| | - | 70 | 29550 | | - | 115 | 47984 |
| Sensibilité = 74,5% [204/(204+70)] Spécificité = 93,4% [29550/(29550+2093)] VPP = 8,9% [204/(204+2093)] VPN = 99,8% [29550/(70+29550)] | | | Sensibilité = 67,8% [242/(242+115)] Spécificité = 97,3% [47984/(47984+1317)] VPP = 15,5% [242/(242+1317)] VPN = 99,8% [47984/(115+47984)] | | | | |

5.3 PROFIL CLINICO-EPIDEMIOLOGIQUE

Une comparaison des profils cliniques et épidémiologiques des cancers dépistés par le programme et des cancers d'intervalle est présentée aux Tableaux 6 et 7, respectivement. En l'absence de données sur l'ensemble des cancers mammaires en Valais dans la tranche d'âge concernée, le profil clinico-épidémiologique des cancers diagnostiqués plus de 2 ans après le dernier dépistage est fourni à titre informatif. Les données cliniques proviennent du RVsT et portent sur les cancers notifiés entre 1999 et 2010 (voir section 5.1).

Le profil pronostique des cancers dépistés apparaît plus favorable que celui des cancers d'intervalle alors que les cancers diagnostiqués hors dépistage chez les participantes plus de 2 ans après leur dernière mammographie ont un profil pronostique intermédiaire. Les différences entre cancers dépistés et d'intervalle sont particulièrement marquées (et statistiquement significatives) pour la proportion d'*in situ*, la différenciation histologique (grade) et le stade agrégé. Tant la taille (taille médiane des cancers dépistés de 14 mm contre 21 mm pour les deux autres groupes de cancer) que la proportion de cancers avec atteinte ganglionnaire sont plus favorables pour les cancers dépistés. Une modeste mais significative différence ($p=0.039$) est également observée dans la répartition anatomique avec une proportion plus élevée de cancers centro-mammaires parmi les cancers d'intervalle. En revanche, la ventilation des types histologiques peut être considérée comme homogène entre les cancers dépistés et les cancers d'intervalle, malgré une proportion légèrement supérieure de cancers lobulaires parmi les cancers d'intervalle.

Sur le plan épidémiologique, on relèvera une proportion significativement plus élevée parmi les cancers d'intervalle de femmes en cours de traitement hormonal ménopausique (48% contre 34% pour les cancers dépistés, Tableau 7). Bien que la proportion de seins denses (BIRADS 3 ou 4) soit plus élevée parmi les cancers d'intervalle que parmi les cancers dépistés (37% contre 31%), on ne peut exclure que cette différence soit le fruit du hasard ($p=0.18$). A noter que les proportions de femmes à seins denses et sous traitement ménopausique sont plus élevées lorsqu'un cancer est diagnostiqué que lorsque la dernière mammographie s'est révélée négative (BIRADS 3-4 : 29,1% ; 24,5% en cours de traitement hormono-substitutif). Enfin, la distribution des cancers dépistés et des cancers d'intervalle selon l'année et l'âge au dépistage indique une sensibilité plus élevée du programme entre 1999 et 2004 qu'entre 2005 et 2008 (72,8% contre 68,8%) ainsi que pour les sexagénaires par rapport aux quinquagénaires (75,2% contre 65,6%).

Tableau 6 Profil clinique des cancers dépistés et d'intervalle du Programme valaisan, 1999-2010

| Caractéristiques cliniques | Cancers dépistés N=448 | Cancers d'intervalle (< 2 ans après dép.) N=185 | Cancers diagnostiqués > 2 ans après dépistage N=281 |
|--|-----------------------------------|---|---|
| In situ* (%) | 13,8 | 5,4 | 10,3 |
| Stade agrégé* (%) | | | |
| Stade I | 60,8 | 39,7 | 47,2 |
| Stade II | 34,0 | 46,6 | 38,1 |
| Stade III ou plus | 5,2 | 13,8 | 14,7 |
| Type morphologique (%) | | | |
| Canalaire | 69,9 | 63,6 | 69,0 |
| Lobulaire | 21,4 | 29,9 | 22,4 |
| Adénocarcinome, autre carcinome | 8,7 | 6,5 | 8,5 |
| Taille* (%) | | | |
| Moins de 10 mm | 21,0 | 20,0 | 15,9 |
| 10-14 mm | 32,2 | 18,2 | 17,5 |
| 15-19 mm | 21,0 | 12,9 | 27,6 |
| 20 mm ou plus | 25,7 | 48,8 | 39,0 |
| Atteinte ganglionnaire* (%) | 27,2 | 36,6 | 36,1 |
| Différenciation histologique* (%) | | | |
| Élevée | 38,5 | 30,4 | 24,7 |
| Moyenne | 47,1 | 44,4 | 56,5 |
| Faible | 14,5 | 25,1 | 18,8 |
| Localisation* (%) | | | |
| Centro-mammaire | 2,5 | 7,6 | 3,9 |
| Interne | 17,9 | 15,7 | 16,7 |
| Externe | 48,2 | 45,4 | 45,6 |
| Quadrants multiples | 25,7 | 23,8 | 27,0 |
| Autre, non renseigné | 5,8 | 7,6 | 6,8 |

* Différence statistiquement significative ($p < 0.05$) pour cette caractéristique entre les cancers dépistés et d'intervalle

Tableau 7 Profil épidémiologique des cancers dépistés et d'intervalle du Programme valaisan, 1999-2010

| Caractéristiques cliniques | Cancers dépistés N=448 | Cancers d'intervalle (< 2 ans après dép.) N=185 | Cancers diagnostiqués > 2 ans après dépistage N=281 |
|---|-----------------------------------|---|---|
| Période de dépistage⁺ (%) | | | |
| 1999-2001 | 19,4 | 20,0 | 22,1 |
| 2002-2004 | 31,3 | 25,9 | 24,6 |
| 2005-2006 | 23,7 | 25,4 | 14,9 |
| 2007-2008 | 25,7 | 28,6 | 38,4 |
| Région (%) | | | |
| Bas-Valais | 37,9 | 36,8 | 34,9 |
| Valais central | 38,4 | 45,9 | 39,5 |
| Haut-Valais | 23,7 | 17,3 | 25,6 |
| Age à la mammographie⁺ (%) | | | |
| 50-54 ans | 21,7 | 26,5 | 17,4 |
| 55-59 ans | 21,4 | 28,1 | 19,9 |
| 60-64 ans | 28,1 | 21,1 | 18,5 |
| 65-70 ans | 28,8 | 24,3 | 44,1 |
| Densité mammaire⁺ (%) | | | |
| BIRADS 1 | 11,4 | 8,1 | 11,7 |
| BIRADS 2 | 58,0 | 54,6 | 55,5 |
| BIRADS 3-4 | 30,6 | 37,3 | 32,7 |
| Traitement hormono-subst.*⁺ (% oui) | 33,8 | 47,9 | 34,1 |

* Différence statistiquement significative ($p < 0.05$) pour cette caractéristique entre les cancers dépistés et d'intervalle

+ Information en provenance du programme de dépistage

Par ailleurs, un examen de l'interprétation radiologique des mammographies précédant le diagnostic de cancer d'intervalle révèle que:

- 27 mammographies présentaient une anomalie positivée par un des deux lecteurs (20 par le R1 et 10 par le R2)^o, soit 14,6% des cancers d'intervalle.
- 14 mammographies positivées après lecture multiple (dont 12 après une L3) ont fait l'objet d'investigations complémentaires qui se sont révélées négatives (7,6% des cancers d'intervalle); 6 de ces investigations étaient invasives.
- Lorsque des clichés comparatifs étaient disponibles (tour incident), le temps écoulé depuis ces clichés n'influence pas le risque de cancer d'intervalle (la sensibilité du programme est comparable indépendamment du délai entre les 2 derniers examens de dépistage).

^o La correspondance entre les lésions positivées et les cancers diagnostiqués ultérieurement ne peut pas être établie.

5.4 COMMENTAIRES

Ces résultats dénotent une progression par rapport à l'évaluation précédente du programme valaisan^p : incidence des cancers d'intervalle en baisse, sensibilité et spécificité en légère hausse. La fréquence des cancers d'intervalle du programme valaisan, environ 2 cas pour 1000 femmes dépistées, est globalement conforme au référentiel européen avec, toutefois, un taux qui demeure marginalement en excès en deuxième année. Le risque de cancer d'intervalle dans l'année suivant la mammographie correspond en Valais à 25% de celui de développer un cancer du sein en population générale et à 50-55% dans la deuxième année (norme recommandée: $\leq 50\%$). Les cancers d'intervalle résultent soit de tumeurs décelables (cancers non vus ou occultes à la mammographie) soit de lésions infracliniques pas encore décelables lors de l'examen et qui ont rapidement progressé. Plus le temps depuis l'examen est court, plus la chance qu'un cancer d'intervalle résulte d'une lésion visible non positivée lors du dernier dépistage est grande. A cet égard, l'incidence contenue des cancers d'intervalle lors de la première année est un signe encourageant de la qualité radiologique du programme.

Relevons que 10 cancers d'intervalle ont été diagnostiqués comme tumeurs *in situ*. L'inclusion de cancers majoritairement issus d'une activité de dépistage hors programme comme cancers d'intervalle du programme est discutable pour diverses raisons, notamment de comparabilité géographique^q. L'exclusion de ces cas rapprocherait l'incidence proportionnelle des cancers d'intervalle en 2^e année des normes européennes (46,3% et 52,7% pour les tours prévalent et incident, respectivement). L'incidence proportionnelle des cancers d'intervalle en 2^e année de la plupart des programmes européens se situe entre 40% et 70%. Ces résultats corroborent ceux d'une étude comparative qui avait rapporté une fréquence relative légèrement plus élevée de cancers d'intervalle en deuxième année en Valais que dans les 2 autres anciens programmes cantonaux (Vaud et Genève)^r.

Les résultats remarquablement similaires en tours prévalent et incident indiquent que le taux de rappel nettement plus bas en tour incident (voir chapitre 3) permet de diminuer le risque de faux positif sans accroître celui de cancer d'intervalle. Avec 7 cancers sur 10 diagnostiqués par dépistage mammographique et 24 examens sur 25 négatifs à juste titre (pas de cancer pendant les 2 années suivantes), les performances globales du programme (sensibilité de 71% et spécificité de 96%) sont comparables aux expériences rapportées ailleurs en Europe^s.

Les caractéristiques des cancers d'intervalle du programme valaisan corroborent largement les observations faites dans d'autres programmes européens avec un profil pronostique clairement plus défavorable que pour les cancers dépistés. L'hypothèse la plus communément avancée pour expliquer cette différence est de nature biologique: les cancers d'intervalle seraient issus d'une population de cancers plus agressifs, à progression rapide, et qui auraient plus de chance de devenir palpables entre 2 mammographies.

Six cancers d'intervalle sur 7 n'avaient été vus par aucun lecteur lors de la plus récente mammographie de dépistage. On ne dispose ainsi pas de description iconographique pour la plupart des cancers d'intervalle, telle que renseignée sur la fiche de lecture pour tous les cancers dépistés. Ces informations, qui permettraient de cerner un profil des lésions potentiellement à risque de devenir des cancers d'intervalle, ne peuvent être obtenues que par une relecture des mammographies de dépistage précédant le diagnostic de cancer. En l'absence de tels renseignements, les éléments identifiés comme pouvant sensibiliser le radiologue à des profils à risque sont les femmes sous traitement hormono-substitutif et à seins denses. Ceci inclut majoritairement les participantes de moins de 60 ans. La sur-représentation des cancers

^p Bulliard J-L, Levi F. Evaluation épidémiologique du programme valaisan de dépistage du cancer du sein: 1999-2009. Lausanne: IUMSP, 2010 (Raisons de santé, no 167).

^q Bulliard J-L, Sasieni P, Klubunde C, De Landtsheer JP, Yankaskas BC, Fracheboud J. Methodological issues in international comparison of interval breast cancers. *Int J Cancer* 2006;119:1158-63.

^r Bulliard J-L, Ducros C, Dayer E, Arzel B, Levi F. Variation in performance in low-volume mammography screening programmes: Experience from Switzerland. *Cancer Epidemiol* 2011;35:293-7.

^s Törnberg S, Kemetli L, Ascunce N, Hofvind S, Anttila A, Seradour B, et al. A pooled analysis of interval cancer rates in six European countries. *Eur J Cancer Prev* 2010;19:87-93.

d'intervalle dans la région centro-mammaire mérite confirmation par d'autres données. Aucun autre programme suisse n'a encore examiné de manière aussi approfondie ses cancers d'intervalle.

Les résultats indiquent qu'au fil des tours la spécificité du programme valaisan augmente et que sa sensibilité baisse (voir par exemple le Tableau 5). Parmi les explications qui pourraient expliquer ce constat, deux méritent qu'on s'y arrête. Tout d'abord, une diminution du recours à l'hormonothérapie - qui augmente la densité mammaire et le risque de cancer d'intervalle - est probable après la publication en 2002 d'une large étude démontrant son effet délétère sur le risque de cancer mammaire. Ensuite, une auto-sélection progressive comme participantes de femmes à moindre risque et dont les cancers sont possiblement plus difficiles à dépister. L'auto-sélection résulte, à chaque tour, de la sortie du programme des participantes chez lesquelles un cancer a été dépisté ou diagnostiqué depuis leur dernière participation, ou chez qui un suivi individuel a été recommandé sur la base du risque personnel. La diminution du réservoir de cancers parmi les femmes fidélisées laisserait ainsi une population de participantes avec des cancers possiblement plus difficiles à identifier par mammographie, d'où une baisse graduelle de sensibilité et une hausse de spécificité du programme. Les résultats des évaluations attendues d'autres programmes romands permettront de valider cette hypothèse.

Globalement, 7,6% des cancers d'intervalle portent sur des mammographies positivées par le programme et sont de probables résultats faux négatifs des investigations complémentaires. Cette proportion se situe à la limite supérieure de celle observée dans les programmes européens (entre 3,5% et 7%).

5.5 RECOMMANDATIONS

La validation des croisements et la comparaison des informations cliniques documentées dans les bases de données du CVD et du RVsT ont mis en évidence certains points perfectibles liés au dispositif d'échange d'informations entre ces 2 bases de données^t. Ces éléments ont ou peuvent avoir des conséquences sur les résultats du programme, notamment la fréquence des cancers d'intervalle. Les plus importants, ainsi que les mesures correctrices proposées, sont résumés dans le Tableau 8.

Tableau 8 Echange d'informations entre le CVD et le RVsT : éléments perfectibles et mesures proposées

| | Observation | Conséquence (potentielle) | Mesure proposée |
|----|---|--|--|
| 1. | Trace des croisements antérieurs non conservée | Sous-estimation des cancers d'intervalle (risque accru de non-identification de certains cas) | Saisie du numéro unique des cas appariés du CVD (resp. du RVsT) dans l'autre base de données* |
| 2. | Informations médicales discordantes | Estimation imprécise de plusieurs indicateurs de performance du programme (si calculé à partir des données du CVD) | CVD: correction (après validation) des paramètres cliniques et thérapeutiques des cancers dépistés (selon la base référentielle du RVsT) |
| 3. | Prévalence non négligeable des faux négatifs des investigations | Augmentation (potentiellement artificielle) des cancers d'intervalle | Examen systématique des faux négatifs |
| 4. | Connaissance lacunaire des cancers d'intervalle | Envois inutiles d'invitations à des femmes inéligibles | Transmission au CVD de tous les cas de cancer du sein diagnostiqués chez des participantes |

* Dû à un prochain changement (2014) de logiciel informatique, le numéro d'identification du CVD va être modifié. Une intégration du futur numéro uniquement est recommandée.

^t Le CVD dispose, à des fins d'évaluation du programme, d'une autorisation fédérale pour obtenir les informations médicales auprès du RVsT pour les participantes consentantes.

La première observation concerne le fichier de transmission des participantes consentantes. Les dossiers du CVD qui ont subi des modifications sur des variables clés pour leur identification (nom de la patiente, date de naissance) ou qui ont été archivés (par exemple découverte de dossiers à double) échappent aisément aux croisements effectués par le RVsT. La conservation, après croisement validé, du numéro unique de la patiente dans l'une (RVsT) ou idéalement dans les 2 bases de données remédierait à ce souci^u. Cela permettrait également de simplifier les futurs appariements, puisque les cas déjà croisés sont aisément identifiables, et de vérifier que tous les dossiers précédemment appariés, ont bien été transmis au RVsT.

Certaines discordances apparaissent également dans la description clinique et thérapeutique des cancers (stadiation, type morphologique, grade, traitement). Le RVsT est la source de référence officielle pour ce type d'information à l'échelle valaisanne. Une identification des différences majeures et une mise à jour des variables cliniques et thérapeutiques renseignées par le CVD pourraient être effectuées à l'occasion de ces croisements. Rappelons que la saisie des mêmes informations par le CVD et le RVsT a été rendue nécessaire afin que le programme dispose rapidement de données cliniques sur ses cancers dépistés. Cette mise à jour requiert une fourniture par le RVsT au CVD de ces variables pour les cancers dépistés (ce n'est pas le cas actuellement).

La troisième observation porte sur les faux négatifs des investigations. Il se peut, vu leur fréquence, que quelques investigations positives, tardives ou non communiquées au CVD, aient servi de base de confirmation d'un cancer. Lorsqu'un tel dossier est « bouclé » comme négatif (cancer exclu) par le CVD sur la base des investigations connues, il est alors de facto considéré comme un cancer d'intervalle. Un examen des cas faux négatifs des investigations est recommandé afin, d'une part, de s'assurer qu'il ne s'agirait pas de sources d'information qui auraient permis au CVD d'enregistrer un cancer dépisté et, d'autre part, d'identifier d'éventuelles filières d'investigations qui auraient conduit à une erreur diagnostique. Ces situations restent, dans l'absolu, rares.

Le dernier élément soulevé dans le Tableau 8 vise à améliorer la qualité de la base de données du CVD mais ne porte pas à conséquence pour l'évaluation épidémiologique des cancers d'intervalle. Présentement, le CVD ne connaît pas systématiquement les cas de cancers du sein diagnostiqués hors programme chez des participantes. Ceci empêche le programme d'arrêter d'inviter des femmes inéligibles (cancer d'intervalle, cancer survenu plus de 2 ans après la dernière participation, cancer avant dépistage mais non communiqué au CVD avant la première participation). Cette situation insatisfaisante pourrait être rectifiée lors des croisements informatiques pour autant que le CVD spécifie au RVsT les dossiers correspondant à des cancers dépistés.

Ces appariements informatiques, indispensables pour identifier les cancers d'intervalle et réaliser une évaluation du programme, sont pour l'instant effectués au rythme des évaluations externes (IUMSP) du programme valaisan. Un croisement annuel des participantes consentantes avec la base du RVsT est souhaitable afin d'assurer une qualité continue et une transmission régulière des cas de cancers d'intervalle au CVD.

^u Cette pratique a par exemple cours au Registre vaudois des tumeurs.

6 CONCLUSIONS ET RECOMMANDATIONS

6.1 SYNTHÈSE

Ce rapport fournit une analyse approfondie et rigoureuse de l'évolution des performances du programme valaisan entre 2002 et 2012, principalement sur le plan de la qualité des lectures, de l'identification et de la quantification des facteurs qui y sont associés, ainsi que des cancers d'intervalle (1999-2010). A ce titre, il s'agit de l'évaluation épidémiologique la plus aboutie d'un programme de dépistage par mammographie en Suisse. Les résultats les plus réjouissants sont:

- Des performances dans l'ensemble conformes aux normes européennes de qualité
- Une tendance générale vers une amélioration modeste et continue des indicateurs principaux de qualité du programme. Ce constat est concordant au travers des divers angles d'analyse (voir chapitres 3 à 5).
- Parmi les améliorations observées, les plus manifestes touchent les performances en 2^e lecture et les quelques indicateurs qui étaient les plus éloignés des normes européennes.
- La fréquence adéquate des cancers d'intervalle, surtout en première année.
- La confirmation des facteurs de risque établis des cancers d'intervalle: un jeune âge (< 60 ans), des seins denses et un traitement hormonal substitutif.
- La coïncidence de l'évolution favorable des performances radiologiques avec deux mesures prises par la direction du programme: un entretien de sensibilisation avec chaque radiologue du programme sur ses performances (en 2006) et l'adoption d'une politique plus stricte d'accréditation des 2^e lecteurs (dès 2007-8). Cette corrélation ne permet cependant pas d'inférer formellement un lien causal entre les actions entreprises et les effets observés.

Deux conséquences importantes et favorables de ces améliorations ont été:

- Une baisse du taux de faux positifs, qui génère des coûts et de l'anxiété chez les participantes, sans augmentation concomitante du risque de cancer d'intervalle.
- Une meilleure répartition des lectures entre radiologues, dans un contexte de croissance du nombre de mammographies, sans affecter le déroulement du programme. En 2012, le volume moyen annuel atteignait pour la première fois 1500 lectures pour les radiologues 2^e lecteurs et 500 pour les 1^e lecteurs.

Les autres résultats saillants de ce rapport sont :

- L'importance du contexte de lecture, du type de tour et, dans une moindre mesure, de l'année de réalisation des clichés comme déterminants de la qualité: les radiologues sont les plus performants lorsqu'ils interprètent des mammographies récentes, en 2^e lecture et chez des femmes qui ont déjà participé (tour incident).
- La grande variabilité du « case-mix » des mammographies interprétées par chaque radiologue.
- Le rôle négligeable de la technologie (analogique ou numérique) et l'effet mineur du type de radiologue (1^e ou 2^e lecteur) et du volume de mammographies lues sur les performances. Ce dernier élément reflète la faible différence de performance entre 1^e et 2^e lecteurs du programme qui ont un différentiel de volume de lectures peu marqué en comparaison avec d'autres programmes.
- Le profil pronostique défavorable des cancers d'intervalle en comparaison avec les cancers dépistés.

- La survenue de la plupart des cancers d'intervalle (6 sur 7) après des mammographies négativées par tous les lecteurs.
- La contribution non-négligeable (env. 8%) des faux négatifs des investigations complémentaires aux cancers d'intervalle du programme.
- La persistance d'un taux de reconvoction (et de faux positif) trop élevé en première participation (8%) avec une proportion également élevée et croissante de cancers dépistés à un stade non invasif (30% d'*in situ* sur la période 2009-11).

6.2 RECOMMANDATIONS

Quelques aspects potentiellement perfectibles du programme ont pu être mis en évidence dans cette évaluation thématique.

En premier lieu, même si les volumes individuels de lecture sont en légère hausse, **le nombre de mammographies interprétées reste modéré (env. 850/an par radiologue, voir Tableau 1)**. En particulier, le nombre de radiologues habilités à faire des 2^e lectures freine l'acquisition d'une expérience solide. Les premiers effets bénéfiques observés des mesures prises sur les performances du programme devraient encourager leur intensification afin d'augmenter le volume individuel de lectures, surtout des 2^e lecteurs. Ceci engendrerait vraisemblablement une disparité plus marquée entre les niveaux de performances des 1^e et 2^e lecteurs, avec possiblement une meilleure complémentarité de la double lecture (meilleure sensibilité du 1^e lecteur conjuguée à la meilleure spécificité du 2^e lecteur). Les possibilités de transmission numérique et le télédiagnostic sont susceptibles de faciliter la distribution des lectures entre radiologues afin d'optimiser les volumes interprétés et tenter de réduire l'effet lié au contexte de lecture.

Les résultats actuels ne permettent pas de recommander de manière tranchée un volume de lectures dans le cadre du programme au-delà duquel un niveau satisfaisant de performances serait assuré. En revanche, ils permettent de déconseiller un seuil annuel en deçà de 1000 mammographies ou un volume cumulé inférieur à 5000 mammographies. Ces seuils, basés sur l'activité dans le cadre du programme, pourraient aussi servir de référence pour l'accréditation de nouveaux radiologues. L'expérience acquise au fil des ans de pratique du dépistage contribue à l'amélioration des performances et devrait être prise en compte dans la considération d'un volume minimal. A cet égard, l'interprétation d'au moins 5'000 mammographies par les 2^e lecteurs est associée à des performances accrues.

Ensuite, le niveau de performance le plus élevé a été systématiquement mesuré en 2006. **Atteindre de manière continue le niveau de qualité observé ponctuellement en 2006 est un objectif réaliste au vu des diverses actions et mesures d'amélioration mises en place depuis lors.** Ceci passe non seulement par la poursuite des mesures de conscientisation des radiologues, avec un feedback annuel sur leurs performances, mais probablement aussi par des entretiens individuels de sensibilisation tels que conduits en 2006 par la direction médicale du programme.

En troisième lieu, **une meilleure compréhension au niveau des radiologues des cancers non vus et des cancers d'intervalle du programme est nécessaire. Ceci permettrait au Programme à la fois d'affiner son feedback aux radiologues^v (p.ex. quel type de lésion est prioritairement ratée) et de mieux apprécier leur niveau individuel de performance.** L'obtention de ces informations requiert une relecture des clichés (du dernier tour pour les cancers d'intervalle et les cancers non vus par un des lecteurs et du tour précédent pour les tumeurs diagnostiquées en tour incident). Une revue des cancers d'intervalle permettrait de catégoriser ces cancers (cancer occulte à la mammographie, cancer non vu, vrai cancer intercalaire, etc.) et de comparer leur distribution avec celle d'autres programmes. **Il est fortement**

^v Ce type d'étude permet par exemple de lier la localisation d'une lésion signalée à celle du cancer diagnostiqué (voir chapitre 5)

conseillé d'inscrire une telle étude uniquement dans le cadre d'un projet de recherche sous supervision épidémiologique, de préférence avec une composante inter-cantonale ou internationale, voire les 2. A cet égard, un projet multicentrique de relecture des cancers (notamment d'intervalle), suivant un protocole épidémiologique strict, est en phase de finalisation au niveau international après 2 ans d'élaboration^w. Dans tous les cas, ces études nécessiteront l'accès à des mammographies numériques (y compris les clichés du tour précédent).

Enfin, une réduction des faux négatifs des investigations diminuerait le taux de cancer d'intervalle du programme. A ce stade, un examen préliminaire de ces cas faux négatifs est fortement recommandé afin d'en déterminer la source (voir aussi section 5.5).

^w Un test international et standardisé de lectures est en phase de finalisation par l'International Cancer Screening Network (ICSN). Tous les programmes membres de la Fédération *swiss cancer screening* seront conviés à y prendre part (relais suisse auprès de l'ICSN : Dr J.-I. Bulliard).

7 ANNEXES

7.1 GLOSSAIRE

Cancer d'intervalle

Cancer du sein diagnostiqué chez une participante dont le résultat de la dernière mammographie de dépistage était négatif. En pratique, seuls les cancers diagnostiqués dans l'intervalle usuel entre deux mammographies successives, soit 2 ans, sont considérés.

Cancer non vu (taux de)

Proportion des mammographies interprétées par un radiologue pour lesquelles la suspicion avérée de cancer est basée sur la conclusion de la lecture d'un autre radiologue (mammographie négativée à tort par le radiologue).

Détection (taux de)

Proportion des femmes dépistées chez qui les investigations complémentaires ont confirmé le diagnostic de cancer (comprend les cancers détectés lors d'un examen de contrôle). Cet indicateur est un estimateur de la sensibilité du dépistage.

Faux négatif (taux de)

Proportion des femmes dont le résultat de la dernière mammographie était négatif et chez qui un cancer d'intervalle a été diagnostiqué. Cet indicateur est un estimateur de la sensibilité du dépistage. Lorsque la femme avait été rappelée mais que le résultat des investigations a infirmé la suspicion de cancer, on parle de faux négatif des investigations.

Faux positif (taux de)

Proportion des femmes rappelées chez qui les investigations complémentaires ont exclu un diagnostic de cancer. Cet indicateur est un estimateur de la spécificité du dépistage.

Incidence en absence de dépistage (taux de)

Taux estimé de l'incidence du cancer du sein dans une population dépistée si aucun dépistage n'avait eu lieu. Cette valeur théorique permet de comparer le taux de détection entre des régions à incidences différentes de cancer du sein. Ce taux est généralement estimé à partir du taux d'incidence pour une période précédant l'activité de dépistage (taux historique) ou sur la base d'une projection statistique (taux modélisé).

Reconvocation (taux de)

Proportion des femmes dépistées qui sont rappelées pour des investigations complémentaires à cause d'une anomalie suspecte (les clichés qui présentent un défaut technique sont généralement refaits sur place et ne sont pas comptés dans cette statistique). Cet indicateur est parfois appelé taux de rappel.

Résultats positifs de la lecture (taux de)

Proportion des mammographies jugées positives par un radiologue.

Sensibilité du programme

Proportion des participantes chez qui un cancer a été dépisté par le programme parmi toutes les participantes diagnostiquées avec un cancer du sein dans les 2 ans suivant l'examen. Cette méthode d'estimation de la sensibilité (dite méthode de détection) dépend de l'exhaustivité de l'enregistrement des cancers d'intervalle et fait l'hypothèse que tous les cancers d'intervalle étaient détectables lors de l'examen de dépistage.

Sensibilité d'une investigation

Proportion d'investigations positives parmi les cancers confirmés.

Spécificité du programme

Proportion des femmes chez qui un cancer a été exclu lors du dépistage parmi toutes les participantes qui n'ont pas été diagnostiquées avec un cancer du sein dans les 2 ans suivant l'examen. Cette estimation de la spécificité (dite méthode de détection) dépend de l'exhaustivité de l'enregistrement des cancers d'intervalle.

Spécificité d'une investigation

Proportion d'investigations négatives parmi les cancers exclus.

Tour de dépistage

Période qui inclut l'ensemble des événements survenant lors d'une séquence d'invitation. D'une durée théorique de 2 ans, un tour débute avec l'invitation de la femme (ou tout autre contact éventuel) et se termine, au plus tard, avec sa prochaine invitation (ou tout autre événement qui justifie de ne plus l'inviter, tel qu'une inéligibilité). Les dates de début et de fin de tour sont propres à chaque femme.

Tour prévalent/incident

Un tour de dépistage est dit prévalent (ou initial) lorsque la femme participe pour la première fois au programme lors de ce tour (même si elle a déjà effectué un dépistage mammographique avant d'entrer dans le programme). Si elle a déjà participé lors d'un tour précédent, le tour est dit incident (ou subséquent), indépendamment du délai écoulé depuis la dernière participation.

Troisième lecture (taux de)

Proportion des mammographies pour lesquelles la discordance entre les 1^e et 2^e lectures nécessite une 3^e lecture. La discordance peut porter sur la notification d'une anomalie par un seul des 2 lecteurs ou sur 2 conclusions positives identifiant des lésions différentes. Dans de rares cas, une 3^e lecture est initiée suite à la fourniture incomplète des résultats d'une des 2 lectures (conclusion, localisation de la lésion).

Valeur prédictive négative (VPN) du test de dépistage

Proportion des femmes non rappelées chez lesquelles aucun cancer d'intervalle n'a été diagnostiqué dans les 2 ans qui ont suivi la dernière mammographie négative.

Valeur prédictive positive (VPP) du test de dépistage

Proportion des femmes rappelées chez qui les investigations complémentaires ont confirmé le diagnostic de cancer (comprend les cancers détectés lors d'un examen de contrôle).

Valeur prédictive positive (VPP) d'un lecteur (ou d'une lecture)

Proportion de femmes rappelées par un lecteur (lors d'une lecture) chez qui un cancer a été dépisté lors du processus de lectures multiples.

7.2 MESURES D'AMELIORATION PRISES

Tableau 9 Description des actions effectuées auprès des radiologues pour améliorer la qualité du Programme valaisan

| Type d'action | Public visé | Période / fréquence | Objectifs | Commentaires |
|---|--|--|---|---|
| Interventions individuelles | | | | |
| <ul style="list-style-type: none"> • Commentaire écrit du Directeur du programme sur les statistiques personnelles • Questionnaire sur les mesures possibles d'amélioration de la qualité • Entretien avec chaque radiologue | Tous les radiologues (accent sur ceux à taux de positifs élevés) | 2003-04 2004 Printemps 2006 | Diminuer le taux de rappel, éventuellement le nombre de radiologues | Ces mesures font suite à 2 rapports d'évaluation indiquant un taux de rappel légèrement trop élevé en tour prévalent |
| Interventions organisationnelles | | | | |
| <ul style="list-style-type: none"> • Organiser auprès des hôpitaux un volume minimal de 500 L1/an • Surveillance des accréditations • Remplacement limité des R2 partants | R1 futur(e)s R1 R1 | 2006-07 / ponctuelle dès 2007-08 / continue dès 2007-08 / continue | Accroître les volumes individuels de lecture, surtout pour les R2 | La restructuration de l'organisation des L1 dans les hôpitaux, lieux où le volume de lectures est le plus bas, n'a pu être réalisée |
| Sensibilisation individuelle | | | | |
| <ul style="list-style-type: none"> • Statistiques des lectures (comparaison d'indicateurs de performances avec les autres radiologues) | Tous les radiologues | dès 2002 / annuelle | Sensibiliser surtout les radiologues dont les performances sont en-deçà de la moyenne | Pas de contact direct avec les radiologues; impact direct sur les performances globales du Programme difficile à estimer |
| Sensibilisation générale | | | | |
| <ul style="list-style-type: none"> • Colloques de lecture | Radiologues participant aux colloques | 2000-08 / irrégulier dès 2009 / annuel | Formation continue | Impact sur les performances globales du Programme difficile à estimer |

L1: 1^e lecture R1: radiologue 1^e lecteur R2: radiologue 2^e lecteur

Source: direction du Programme valaisan de dépistage du cancer du sein, mars 2013.

7.3 RESULTATS DES ANALYSES STATISTIQUES DE TENDANCES DES PRINCIPAUX INDICATEURS DE QUALITE

Figure 9 Tendances annuelles (2002-11) des principaux indicateurs de qualité du Programme valaisan par tour de dépistage (analyse Joinpoint)

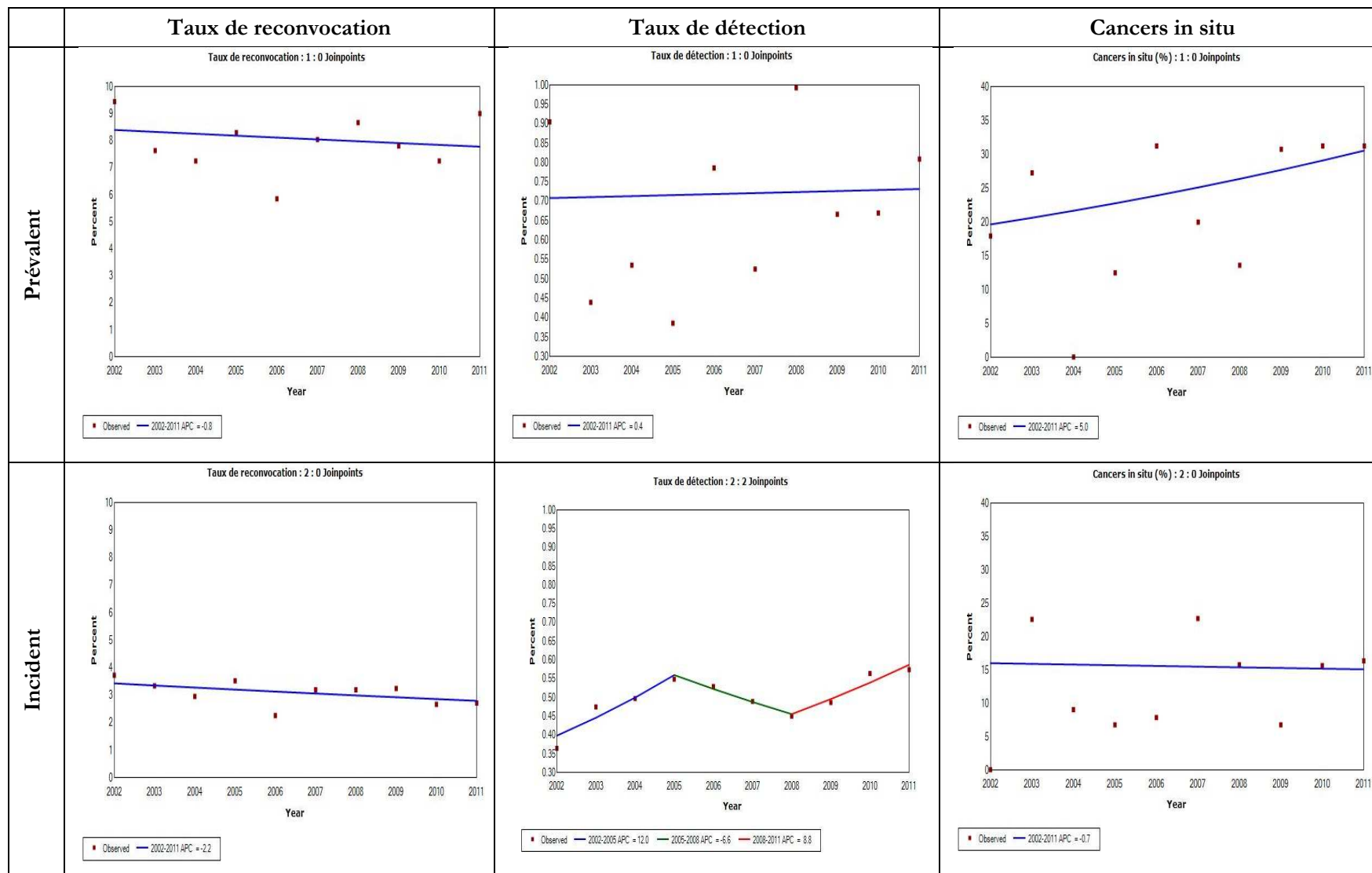


Figure 9 (cont.) Tendances annuelles (2002-11) des principaux indicateurs de qualité du Programme valaisan par tour de dépistage (analyse Joinpoint)

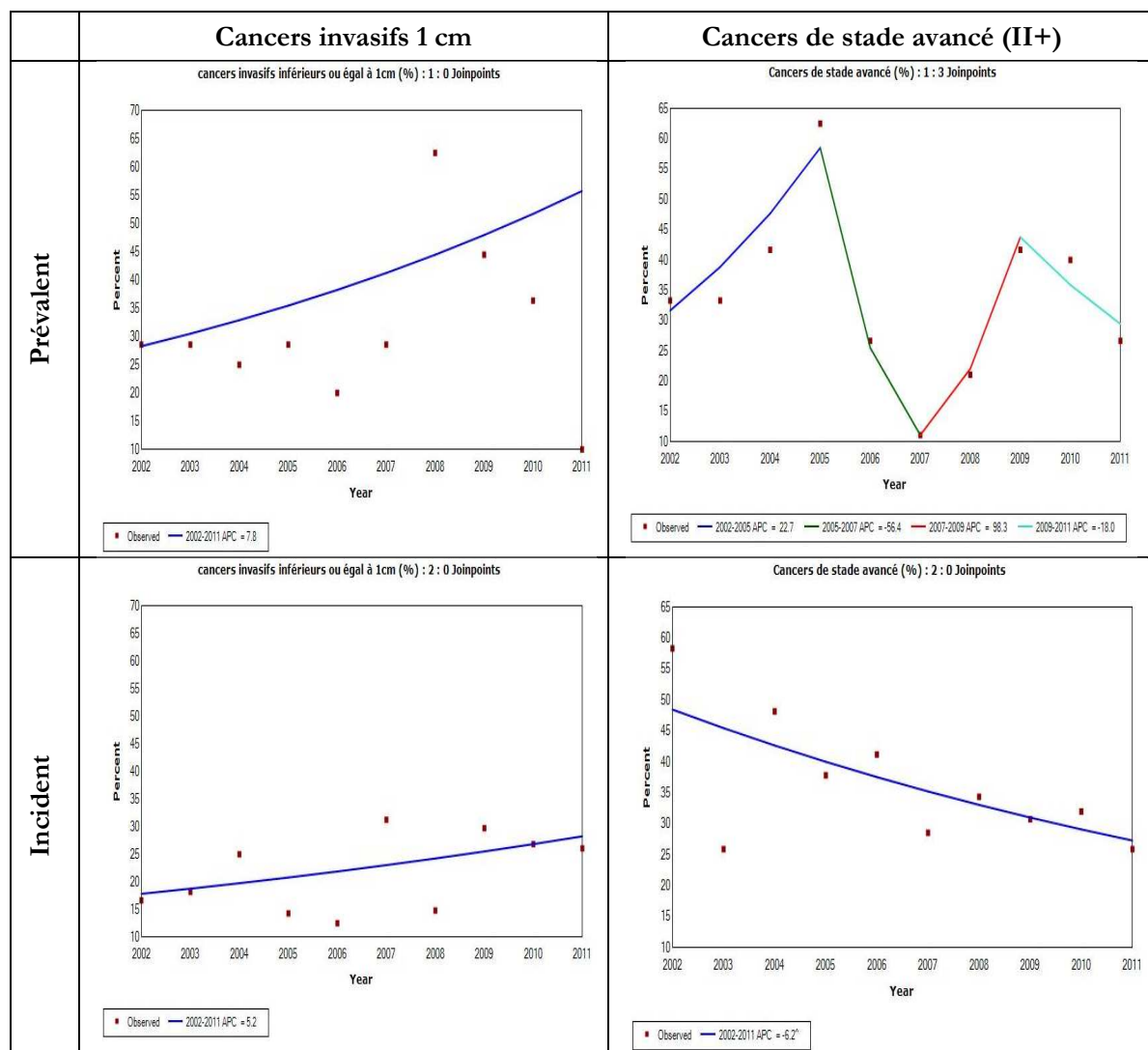


Figure 10 Tendances annuelles de la valeur prédictive positive en 1^e et 2^e lecture du Programme valaisan (analyse Joinpoint)

