

## Introduction

### Contexte

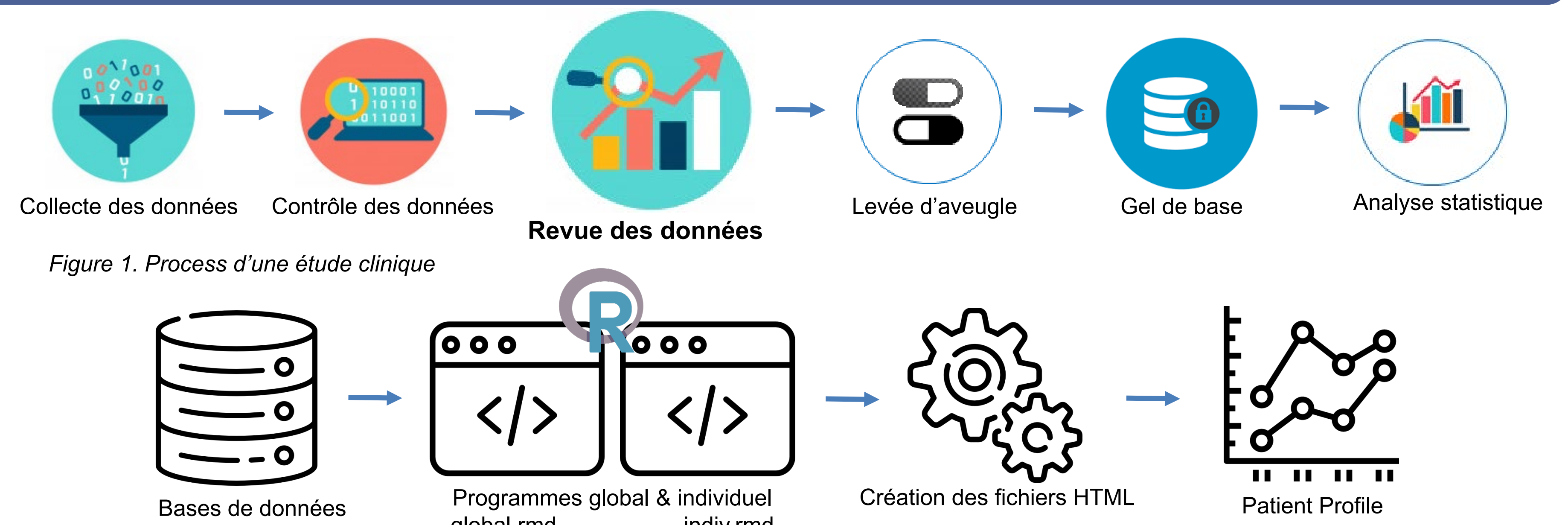
Dans les essais cliniques, une étape importante est la **revue des données** en aveugle (Figure 1). Celle-ci intervient entre la fin de l'essai (quand les données de l'étude sont collectées, saisies et validées) et le gel de base puis la levée d'aveugle. L'objectif est de revoir les données et les conditions de réalisation de l'étude (déviations au protocole, événements indésirables, traitements concomitants, valeurs extrêmes...) qui peuvent avoir une incidence sur les résultats, afin de prendre des décisions sur la validité des données ou des sujets à exclure de certaines populations d'analyses.

### Objectif

Afin de faciliter la réalisation de cette phase des projets cliniques, les **données importantes peuvent être présentées par sujet**, incluant certaines caractéristiques à l'inclusion (âge, sexe, antécédents médicaux...) et des informations sur les critères de jugement et le suivi (paramètres anthropométriques, biologiques, déviation au protocole, compliance...). Ces informations sur les profils des patients, regroupées au sein d'un **outil que nous appelons 'patient profiles'**, étaient jusqu'à présent disponibles en format PDF. Ce document, créé à l'aide du logiciel SAS, était statique et contenait de nombreuses pages, limitant sa maniabilité. Notre objectif était d'améliorer la revue des données avec un **document interactif afin de simplifier la navigation et la liaison entre les différents types de données d'un même sujet**.

### Méthode

Un outil permettant de produire les *patient profiles* au **format HTML via R Markdown** a été développé (Figure 2). R Markdown est un format permettant de créer des rapports dynamiques et reproductibles avec R (tableaux et graphiques interactifs, liens hypertextes, navigation entre onglets...). Le code R Markdown génère des fichiers HTML qui contiennent les résultats (tableaux et graphiques) pour chaque sujet et qui peuvent être consultés à partir de n'importe quel navigateur par les personnes impliquées dans la revue des données. Ce travail a été réalisé avec la version 4.0.2 de R et l'utilisation de plusieurs packages tels que markdown, haven, dplyr, shiny...



### Résultats

Plusieurs fichiers sont créés. Un fichier global ainsi qu'un fichier par sujet. Le fichier global contient un onglet « Information » avec la liste des tables utilisées, une barre de recherche ainsi que la liste de tous les sujets pour permettre d'y accéder rapidement. Il contient également un 2<sup>ème</sup> onglet « Data Overview », regroupant quelques **graphiques réalisés sur l'ensemble de la population** tels que des boîtes à moustache et des courbes individuelles (Figure 3).

Les fichiers individuels contiennent à gauche plusieurs onglets rassemblant les informations générales (âge, sexe, centre investigateur, caractéristiques du sujet), antécédents médicaux, dates de visites, compliance, déviations au protocole et examen clinique. Ces différentes informations sont affichées sous forme de **tableaux ou de listes**. Enfin, la partie à droite contient différents **graphiques interactifs** permettant d'étudier l'évolution des événements liés à un sujet (événements indésirables, traitements concomitants, visites), ses paramètres anthropométriques, cliniques et biologiques ainsi que les valeurs de la population (minimum, maximum et quartiles). Un survol de la souris sur les courbes permet d'afficher des informations additionnelles comme la valeur observée, le numéro du sujet et la visite (Figure 4).

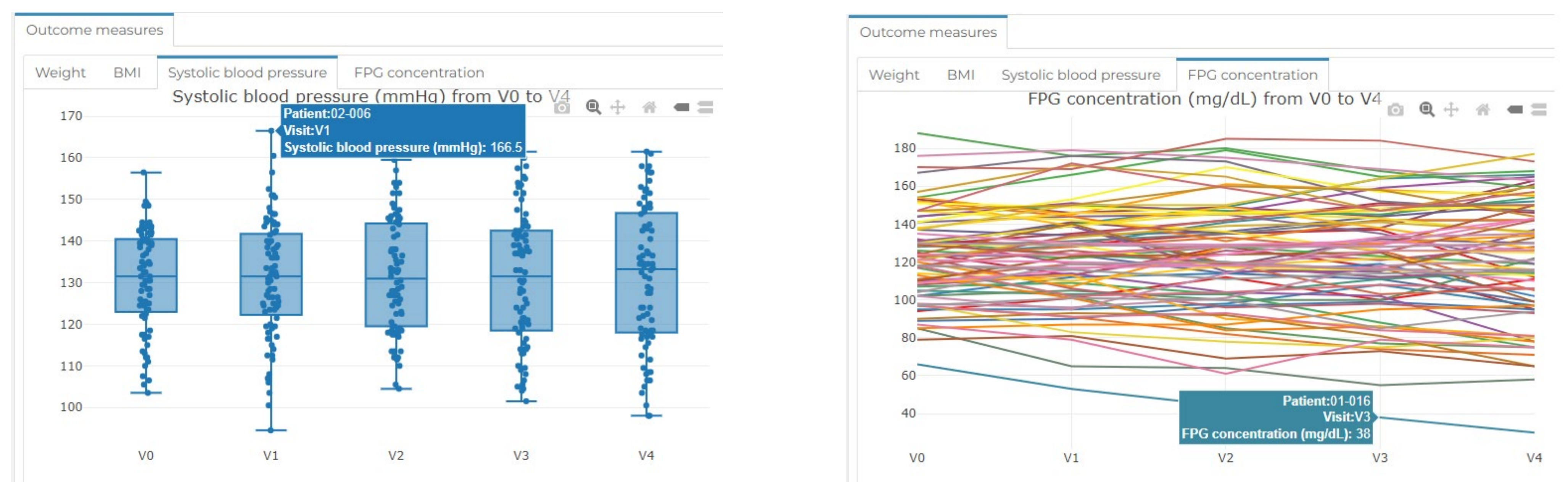


Figure 3. Graphiques sur l'ensemble des sujets de la pression artérielle systolique et de la concentration plasmatique à jeun du glucose

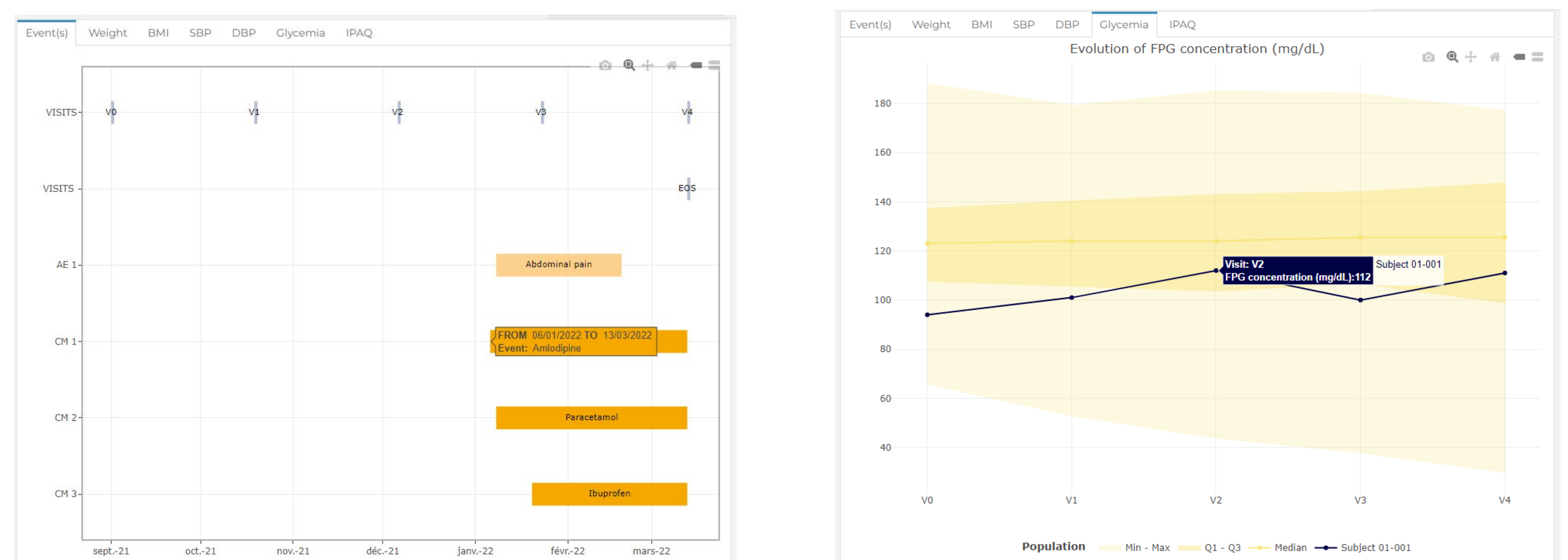


Figure 4. Graphiques individuels comprenant les événements du sujet (visites, événements indésirables, traitements) et l'évolution de sa glycémie

### Conclusion

Les *patient profiles* codés en HTML ont un **visuel beaucoup plus interactif et ergonomique** que ceux réalisés précédemment au format PDF. Ils permettent de naviguer rapidement au sein des informations d'un sujet et entre ceux-ci afin d'**étudier la qualité des données** (détection des données aberrantes ou des déviations), **d'associer certains paramètres** (événements indésirables et traitements concomitants par exemple) et ainsi **préparer au mieux l'analyse statistique ultérieure**.

### Références

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.  
JJ Allaire and Yihui Xie and Jonathan McPherson and Javier Luraschi and Kevin Ushey and Aron Atkins and Hadley Wickham and Joe Cheng and Winston Chang and Richard Iannone (2021). rmarkdown: Dynamic Documents for R. R package version 2.11.  
Yihui Xie and JJ Allaire and Garrett Grolemund (2018). R Markdown: The Definitive Guide. Chapman and Hall/CRC. ISBN 9781138359338.

### Remerciements

Merci à Valentin Redon et Mélanie Demeure qui ont contribué à améliorer grandement l'outil des *patient profiles*.