



This is a repository copy of *Large scale studies assessing anti-factor VIII antibody development in previously untreated haemophilia A: what has been learned, what to believe and how to learn more.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/110421/>

Version: Accepted Version

Article:

Iorio, A., Fischer, K. and Makris, M. orcid.org/0000-0001-7622-7939 (2017) Large scale studies assessing anti-factor VIII antibody development in previously untreated haemophilia A: what has been learned, what to believe and how to learn more. British Journal of Haematology, 178 (1). pp. 20-31. ISSN 0007-1048

<https://doi.org/10.1111/bjh.14610>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Large scale studies assessing anti-factor VIII antibody development in previously untreated haemophilia A: what has been learned, what to believe and how to learn more.

Alfonso Iorio (1), Kathelijn Fischer (2), Michael Makris (3,4)

- (1) Department of Clinical Epidemiology and Biostatistics and Department of Medicine,
McMaster University, Hamilton, Canada
- (2) Van Creveldkliniek, University Medical Centre Utrecht, Utrecht, The Netherlands
- (3) Sheffield Haemophilia and Thrombosis Centre, Royal Hallamshire Hospital,
Sheffield, , UK
- (4) Department of Infection, Immunity and Cardiovascular Disease, University of
Sheffield, Sheffield, UK

Word count: 6184

Abstract word count: 204

Running title: Differential immunogenicity in haemophilia A PUPs

Keywords: haemophilia, inhibitors, epidemiology, immunogenicity, FVIII concentrate,
Previously Untreated Patients

Address correspondence:

Alfonso Iorio
Clinical Epidemiology & Biostatistics <<http://fhs.mcmaster.ca/ceb/>>

McMaster University
1280 Main St West
Hamilton, ON L8S 4K1
Tel: +1 905 525 9140 ext 22421
Fax: +1 905-526-8447
e-mail: iorioa@mcmaster.ca

Summary

Minimizing the risk of inhibitor development by acting on modifiable risk factors remains a sensible goal for treatment optimization in haemophilia A. By critically appraising published studies assessing inhibitor development, we address the role of studies in previously untreated patients (PUP) for establishing the immunogenicity of new concentrates, suggest novel research design to be adopted in future studies, and discuss clinical practice implications of the reported differential immunogenicity of Kogenate and Advate factor VIII concentrates. Three considerations are relevant here: i) all of the existing concentrates, when tested following the ISTH SSC recommendation, were shown to be safe; as a consequence, ii) when considering using any newly introduced product, one should be aware that it could in the future turn out to be as immunogenic as Kogenate iii) at the population level, it might be wiser not to use Kogenate in PUPs, if the choice is against Advate. When presenting the risk of developing inhibitors to the individual patient (or to his family), the message remains that the risk can be as high as 40%, without any efficient instrument to predict individual inhibitor risk. Patients should be invited to enrol into a randomized registry trial, including random assignment to trials with new investigational products.

Background

Hemophilia A is a congenital bleeding disorder with genetically determined absence or reduction of clotting factor VIII. Persons living with haemophilia experience lifelong spontaneous bleeds in their joint, which trigger an inflammatory response leading to haemophilic arthropathy (Mannucci & Tuddenham, 2001). FVIII concentrates are the mainstay of treatment of severe and moderate haemophilia A, particularly when given regularly to prevent bleeding (Marchesini *et al*, 2011). The introduction of viral inactivation of concentrates in the mid 1980s has virtually eliminated the risk of HIV, HCV and other infections. Although concentrates are now safer than they have ever been, the development of inhibitors (alloantibodies to FVIII) occurring in up to 30% of previously untreated patients (PUPs) remains the single most important obstacle to haemophilia management at present.

The development of inhibitors has recently been the subject of intense investigation using different approaches. Basic research has made enormous progress towards identifying key mechanisms that can be modulated to prevent and treat inhibitors, which in the future could dramatically change treatment opportunities (Matino *et al*, 2015; Sack *et al*, 2014; Scott, 2014a, 2014b; Gupta *et al*, 2015; Wang *et al*, 2015; Dolgin, 2014). Until then, minimizing the risk of inhibitor development by acting on modifiable risk factors will remain the mainstay of treatment optimisation. In this perspective, a lot of innovative epidemiological evidence has been recently generated (Gouw *et al*, 2013a, 2013b; Calvez *et al*, 2014; Collins *et al*, 2014; Fischer *et al*, 2015; Marcucci *et al*, 2015; Peyvandi *et al*, 2016). Indeed, since 2013, several large scale epidemiological studies and a randomised controlled trial have been published, largely focusing on or being influenced by the differential immunogenicity of specific factor concentrates. These studies constitute significant progress in inhibitor knowledge, which goes beyond the comparison of immunogenicity of different concentrates.

The scope of this manuscript is to critically review the recently published studies assessing the development of inhibitors in haemophilia, with the specific objectives of:

- i) assessing the value and limitations of the recently published studies
- ii) addressing the role of PUP studies for assessment of immunogenicity of concentrates in the future
- iii) suggesting novel research design to be adopted in future studies
- iv) discussing the implication for clinical practice of the available evidence for differential immunogenicity of different factor VIII products.

Strengths and limitations of the recently published studies

A synopsis of the study characteristics and results of the studies published since 2013 is provided as Tables 1 and 2. In general, all papers reported on all titer inhibitors as main analysis, and adjusted the analysis for some of the following covariates: ethnicity, F8 gene mutation, disease severity, family history of hemophilia and inhibitors, history of severe bleeding, age at first exposure, previous exposure to blood components, reason of first treatment, dose of FVIII, FVIII source, peak treatment moments, history of

switching between product brands, major surgery and regular prophylaxis, regular prophylaxis initiation, duration between exposure days, birth year, calendar period, country, and treatment center). There was overlap in the enrolled population among the studies we have reviewed, which have been accounted for in the analysis presented in table 2.

There is a wide unexplained difference in average (baseline) inhibitor risk among the various cohorts, larger for risk of overall inhibitor (spanning from 21.2 to 33.3%, I^2 for heterogeneity 64.7%) than for risk of high titer inhibitors (ranging from 14.6 to 20.2%, I^2 for heterogeneity 46.5%). Particularly high heterogeneity was found across the studies for subgroups of patients treated with recombinant products as a class and for patients treated with Kogenate (Table 1), whereas less was found for plasma derived concentrates or Advate. One possible interpretation of this is that plasma derived concentrates and Advate performed more similarly across the studies than Kogenate did. Overall, a mixed model analysis showed statistically significant differences for the comparison of plasma derived versus recombinant factors for all inhibitors and for Kogenate versus Advate for all and high risk inhibitors (details in the legend of Table 1). Part of the difference can be the effect of the variability in the definitions of inhibitors adopted in the different studies and the different strategies used to assemble the various cohorts, though all are reported as being inception cohorts (see table 2 and Supplementary Table S1 and S2 for details). Whilst this might not be a problem for within-cohort comparisons, it might inflate the variability of across cohort and pooled analyses. In addition, some studies span a long observation period and very different settings, therefore include differences in treatment strategies between centres as well as differences in treatment intensity over time (Nijdam *et al*, 2015).

The RODIN project and reports

The RODIN study report (Gouw *et al*, 2013b) explored immunogenicity of factor concentrates in PUPs treated in 29 centres throughout Europe, Canada and Israel. The authors reported on 177 inhibitors (all titres) that developed in 574 (31%) haemophilia A PUPs born 2000-2010 and who were observed until inhibitor development or reaching 75 exposure days (ED). The pre-specified hypothesis of the study(Fischer *et al*, 2014a) was of a differential immunogenicity between plasma derived and recombinant concentrates considered as separate classes; this hypothesis was rejected. However, this study observed only 29 inhibitors in 88 patients treated with plasma derived concentrates, which is likely too small a group for a sufficiently powered comparison. However, using the patients treated with Advate as a reference group, this paper showed a higher rate of inhibitors in patients treated with Kogenate/Bayer/Helixate NextGen. The comparison in this paper has been criticized for its choice of reference group; if each of the recombinant products had been compared to the plasma derived group, no significant difference would have been observed. Similarly, if the authors had taken the natural approach of choosing as reference group the largest one (Kogenate/Helixate NextGen, with 183 patients), they would have concluded that, likely by chance, the 157 patients treated with Advate had reported a statistically significant lower number of inhibitors at multivariable analysis. Interestingly, the companion paper reporting on the role of prophylaxis in inhibitor development (Gouw *et al*, 2013a) did not adjust for the generation of recombinant

products, but follows the original analysis plan adjusting for the difference among classes of concentrates, though this was found not significant in the original publication (Gouw *et al*, 2013b). This paper did not report details such as breakdown of exposure and inhibitor development by centre, regression coefficients for the covariates used in the adjustment or temporal trends in use of the various concentrates and in the rate of inhibitors. However, the PedNet group unequivocally showed the feasibility of multi-sponsored, large scale, high accuracy data collection on the inhibitor development in PUPs, producing a mass of data sufficient to enable comparisons between brands. In this perspective, RODIN has paved the way to prospective comparisons of immunogenicity of concentrates in PUPs.

The UKHCDO database report

The UK Haemophilia Doctors Organisation (UKHCDO) has since 1968 maintained a national registry, which has been used to generate several research reports (Darby *et al*, 1996, 2007, 2004; Hay, 1998; Hay *et al*, 2011; Colvin *et al*, 1995; Björkman *et al*, 2010). Following the publication of the RODIN study, Collins et al used this registry to compare the immunogenicity of all recombinant FVIII products available in the UK market (Collins *et al*, 2014). Between 2000 and 2011, 118 inhibitors developed in 407 severe haemophilia A PUPs (29%). Significantly more patients developed inhibitors on Kogenate Bayer/Helixate NextGen than on Advate, both at unadjusted and at adjusted analyses. Since 5 large UK centres participated in the RODIN study, there was a 22% overlap between the studies. Rather surprisingly, the rate of inhibitor development in patients treated with Kogenate Bayer/Helixate NextGen in the UK-RODIN centres was borderline significantly higher than in non-RODIN centres, theoretically supporting the alternative hypothesis of a “RODIN-centre effect”. Why should the inhibitor rate for Kogenate Bayer/Helixate NextGen but not Advate, be higher in the RODIN centres? Another interesting observation in the UKHCDO cohort stems from data on exposure days being incomplete. To accommodate for the missing information the authors hypothesised that for patients on regular prophylaxis calendar time can be used to measure exposure instead of exposure days. In practice, the authors used data from the UK-RODIN centres, which had measured exposure days in all patients, to estimate the average time needed to reach 50 ED, and used it to analyse the rest of the cohort. If confirmed to be valid in other cohorts, using calendar time instead of ED could make future studies much easier to perform and analyse. (Iorio *et al*, 2012). Of course, ED and time on treatment may be strongly correlated for patients adherent to prophylaxis, but not for patient not adherent to prophylaxis or treated on demand, for whom recording and analysing both ED and time might be needed. One point of strength of the UKHCDO cohort is that the choice of concentrates in UK is largely driven by a tender process, which would reduce the likelihood of selection bias. Indeed, if the choice of treatment is driven by the availability or not of a specific product, one may analyse the data under the assumption of the “paired availability approach” (Baker *et al*, 2001; Baker & Lindeman, 1994), comparing rates of events before and after the tender; this is considered one of the more robust among the observational designs.

It must be noted that all these studies spanned over a decade, with variation in the proportion of patients treated with one or the other product over time. The UKHCDO

cohort shows that the relative rate of the immunogenicity of Kogenate Bayer/Helixate NextGen and Advate has changed over time (Figure 1), which is likely due to some yet unknown risk factors (unmeasured confounding). Many hypotheses can be proposed to explain this variability over time. We would suggest considering two of them: first, selection bias, which could have acted through selection of high or low risk patients for different products depending on current beliefs about differential immunogenicity; second, some transient modification of the manufacturing process for Kogenate Bayer/Helixate NextGen.

Finally, the UKHCDO analysis also suggested a signal for higher immunogenicity of the B-domain deleted product. Whilst the power for this comparison is lower, this finding provides further support to the feasibility of comparing immunogenicity in PUPs. The UKHCDO cohort included no plasma derived concentrates, as the UK national guidelines recommend the use of recombinant concentrates for PUPs following the demonstration of transmission of variant CJD by blood products (Keeling *et al*, 2008).

The FranceCoag database report

A further effort to confirm or refute the RODIN result has been proposed by the FranceCoag group (Calvez *et al*, 2014). Though based on a nationwide comprehensive inception cohort maintained since 1994, only 303 of the total 741 patients were included in the study. Among the patients excluded, were 50 enrolled in the RODIN study and 110 who received plasma derived concentrates at the first infusion. The overall inhibitor rate reported is 38% (114/303), which is higher than in the other studies. Thirty-three of 97 (34%) patients treated with Advate and 55 of 111 (50%) patients treated with Kogenate Bayer/Helixate NextGen developed an inhibitor. The difference, though impressive, was not statistically significant both as proportion and as unadjusted or adjusted hazard ratio. Of note, the absolute rate of inhibitors with Advate is significantly higher than that observed in RODIN and particularly in the UK, where plasma derived concentrates were not an available alternative. A recent communication from the same group (Goudemand *et al*, 2015) has reported a significant reduction in the risk of developing inhibitors when comparing the patients treated with Factane (20 inhibitors in 99 patients, crude rate 20.2%) with patients treated with Advate (37 inhibitors in 121 patients, crude rate 30.2%). The hazard ratio at multivariable analysis was found to be 0.53 (95% CI 0.29-0.98, p 0.042). Interestingly, if we add these data (Goudemand *et al*, 2015) back to the original publication (Calvez *et al*, 2014), the overall inhibitor incidence would be lower at 33% (134/402).

The French investigators have to be congratulated for the level of detail provided in their comprehensive supplementary analysis. One of their analysis is the one-out sensitivity analysis (a set of meta-analyses where the data were re-analysed many times, each one time leaving one center out). Based on this analysis, and on some theoretical background, Berntorp and Iorio have suggested how variability among centres could influence the overall results (Berntorp & Iorio, 2015). Clearly the analysis does not provide an alternative explanation, is post-hoc, and not adjusted for the many covariates, but is suggestive. So far, the only study publishing inhibitor development according to concentrate whilst adjusting by centre has been the UKHCDO report. Recently, RODIN reported a sub-analysis suggesting the absence of centre effect (Van den Berg *et al*,

2015). However, these results cannot be compared to the UKHCDO data because they only included a comparison of findings in larger versus smaller centres. This sub-analysis showed that the larger centres had higher inhibitor rate (34% vs 24%), mostly due to more low titer inhibitors detected with more frequent testing (5 vs 3 tests per patient respectively) and the use of higher average weekly dose of factor VIII concentrate (82 vs 68 IU kg⁻¹ ED⁻¹ for the first 5 EDs). As to the comparison between Kogenate/Helixate NextGen and Advate, the authors only reported that there was no difference in the proportion of usage of Kogenate Bayer/Helixate NextGen in the two categories of centres. These analyses of the effect of centre size strengthen the hypothesis that a centre effect might have played a role in the RODIN study. Analyzing for a centre effect as proxy of unmeasured confounding would require a different analytical approach (Hougaard, 1995; McGilchrist & Aisbett, 1991).

Another interesting consideration proposed by the French authors to help understand the objective dimension of the increased immunogenicity of Kogenate/Helixate NextGen is the estimation of the annual number of new inhibitors that would be observed in France due to the increased risk with Kogenate/Helixate NextGen. This number would be 1 to 2 per year (Calvez *et al*, 2014), which is not negligible, but is likely to be less than one would expect as the result of observing an HR of 1.6 for Kogenate/Helixate NextGen over Advate. This would also imply that if one to two inhibitors per year on Kogenate Bayer/Helixate NextGen in the FranceCoag cohort were actually explained by an underlying independent cause (e.g. confounding by indication), the results of the study would be reversed.

The EUHASS surveillance system

Recently, the EUHASS group has reported a comparison of the inhibitor rates with different concentrates in the first 4 years of their data collection in 57/60 participating centres (Fischer *et al*, 2015). Designed as a safety surveillance system, EUHASS is a large study, with a very light protocol, which facilitates uptake by many centres. The way the data are collected in EUHASS is different from the previous reports in several critical aspects: firstly, participating centres are requested to provide data for all their patients (inception cohort); secondly all data on inhibitors are prospectively collected; thirdly, while the exact ED and risk factor are collected for the cases, such data are not available for the non-cases, so adjusted Cox regression is not possible; fourthly there is only limited data validation, no follow up data and no information on peak titers, so that the distinction between high and low titers is less clear (Makris *et al*, 2011; Fischer *et al*, 2011). EUHASS has recruited in 4 years almost the same population enrolled in more than double the time in RODIN – on the other hand, the only benefits EUHASS can claim are that of a large sample size and products, where ideally random difference will be the only one observed. EUHASS did not find a statistically significant difference in inhibitor development according to concentrates in the overall population, but in the subset of patients uniquely reported to EUHASS (Fischer *et al*, 2016), a trend for a higher rate in Kogenate Bayer/Helixate NextGen (RR 1.09, 95% CI 0.58 – 2.04 for all inhibitors and 1.57, 95% 0.68 – 3.60 for high titre) was observed. The non-significance in EUHASS has to be appraised against the smaller sample size and the unadjusted statistical analysis. Assessing significance via non-overlapping confidence intervals of rates is much less

powerful than multivariable Cox regression to assess time to event; on the other hand, the EUHASS analysis approach is easier for clinical practice application, and shows directly that the difference, if any, is small (as shown in the post-hoc simulation of the French cohort). On theoretical grounds, the EUHASS is the single study trying to prospectively answer the specific question of differential immunogenicity of different concentrates. It has to be acknowledged, however, that the superiority of the pre-specified hypothesis can become a source of bias in unblinded studies (like EUHASS), in that patients can be selected for specific products, or even are more likely to be observed and reported on the base of external evidence. For example, more inhibitors could have been detected and reported for Kogenate Bayer/Helixate NextGen after publication of the RODIN study. However, this bias can only occur in future studies, as EUHASS included data collected up to December 2012, before the RODIN publication.

The EAHAD IPD meta-analysis

This study (Marcucci *et al*, 2015) was performed under the auspices of the European Association for Hemophilia and Allied Disorders (EAHAD). For this meta-analysis the authors assembled a cohort of 761 PUPs, a subgroup of whom provided evidence for the comparison between Kogenate Bayer/Helixate NextGen and Advate. The results are very similar to those of the EUHASS report, with a non-significant signal of a moderate increase in risk. Limitations of this analysis are the extremely long time span of the data collection (from 1935 to 2010), including many changes in treatment strategies, the lesser richness of covariate data and some of the assumptions made (e.g. attributing any observed inhibitor to the first concentrate the patient was treated with), the retrospective data collection, incomplete data verification, heterogeneity and incompleteness of data collection. However, the approach presents several valuable points: first of all, it is an example of shared databases, where different authors share their original data for a common and independent analysis. Second, it presents two alternative approaches to the analysis of inhibitor development data, namely the use of propensity scores and classification and regression tree (CART) analysis. Interestingly, the CART analysis suggested that recombinant FVIII concentrate choice was not an important determinant of inhibitor development. The adoption of different statistical approaches, based on different assumptions, might be useful to increase our confidence in the direction and size of the observed effect.

The EMA PRAC meta-analysis

Following the RODIN, UKHCDO and FranceCoag publications the EMA has collaborated with the coordinators of these studies and performed an individual patient meta-analysis. A recent press release by the EMA PRAC (EMA/PRAC/332348/2016, 2016) showed the same overall effects we showed, but concluded that too much variability was left unexplained and too much confounding was still possible, thus making it impossible to draw firm conclusions. However, the detailed results from the final analysis are awaited.

The SIPPET study

The Survey of Inhibitors in Plasma-Product Exposed Toddlers (SIPPET) was a randomized study comparing immunogenicity of plasma derived and recombinant factor concentrates in 251 PUPs or minimally treated patients followed up for 50 ED or 3 years (Peyvandi *et al*, 2016). The study considered different plasma derived and different recombinant products as equivalent to each other, aiming to test the existence of a class effect. A post-hoc analysis was introduced after publication of the RODIN study results. The study design and analysis present some threats to its internal validity, including knowledge of the arm the next patient would be assigned to (inadequate concealment of allocation), premature termination of an open study, incomplete follow up, and major deviations from the published protocol (including a different statistical analysis plan, and a change in the definition of clinically relevant difference). Concerns to the external validity of the results have been also raised (MASAC; Fischer & Blatny, 2016; van den Berg *et al*, 2016; Iorio, 2016). First, the rate of inhibitors, both overall and high responding, is higher than usually observed (Table 1), and particularly so in the recombinant group (37% for all inhibitors, 24% for HR, of which 93% persistent). One or more of the following can explain this high rate: most prevalent ethnicity of the assembled population, very low usage of prophylaxis, and selection of very high-risk patients. Altogether, all these characteristics of the SIPPET population making it very difficult to directly apply the observed difference in immunogenicity to patients in the Western world. This point is particularly relevant when calculating number needed to treat (NNT) to avoid one inhibitor case. Since the NNT depends on prevalence, the NNT in a population with lower risk of inhibitors would need to be recalculated and would be much larger. As to the decision to stop SIPPET early, with the consequent risk of overestimating the true effect in a trial with low absolute number of events (Bassler *et al*, 2010), it will be interesting to compare the SIPPET results with a still unpublished analysis of the FranceCoag cohort. The French author reported a statistically significant higher rate of inhibitors for Advate versus the LFB plasma FVIII concentrate at the ISTH in Toronto, on July 2015 (Goudemand *et al*, 2015), and effect which disappeared after 6 additional months of follow up, as reported at the EAHAD meeting in Malmo, February 2016. This is exactly what is expected when stopping trials early (Guyatt *et al*, 2012). We recommend waiting for the final results of the FranceCoag results before drawing any conclusion on the applicability and clinical relevance of the SIPPET study.

Moving to the comparison between Kogenate and Advate, there are two aspects of the SIPPET study that deserve mention. Thought the SIPPET results are not reported in sufficient detail to allow calculation of the rate of inhibitors for patients on Kogenate, the exclusion of the centres using Kogenate drives the OR for the risk in recombinant treated patients up from 1.87 (95% CI, 1.17 to 2.96) for all inhibitors and 1.69 (95% CI, 0.96 to 2.98) for HR inhibitors to 1.98 for all inhibitors (95% CI, 0.99 to 3.97) and 2.59 (95% CI, 1.11 to 6.00) for high-titer inhibitors. This might indicate that either the risk in plasma derived treated patients was lower in the recombinant arm of centres using Kogenate, or the risk on Recombinate (which was used in the vast majority of the remaining patients) was higher than on Kogenate. In both cases, the SIPPET study, though not directly aimed at addressing the question of differential immunogenicity of recombinant concentrates, would not support the case for a higher rate of inhibitors in Kogenate treated patients.

These data would also suggest that the class effect in plasma derived concentrates might not be confirmed by SIPPET. Unfortunately, these results have not been included in the NEJM publication, and have not been disclosed to date.

A final point that was not addressed by the study but has to be taken into consideration is that, though shown safe in practice, plasma derived factor concentrates still have a higher theoretical risk of transmission of emerging blood borne infections, either virus or prion related.

General considerations

Taken together, the six studies reviewed outline the feasibility and power of large data collection and experimental design in terms of hypothesis generation and potential impact on the public and regulators. However, there is clearly a need for a precise research agenda to overcome current limitations and pave the way to better use of existing and future data. The single and most important limitation of all the observational studies included is that there is still a high risk of selection bias, i.e. different products being chosen for patients at different inhibitor risk.

Sharing data is becoming the standard of practice in many areas of medicine (Drazen, 2015; Lo, 2015), thanks to the “All trials” campaign (<http://www.alltrials.net/>), the NHS (MRC *et al*, 2015) and the Institute of Medicine (Committee on Strategies for Responsible Sharing of Clinical Trial Data Board on Health Sciences Policy Institute of Medicine, 2014). Unfortunately, haemophilia is no better than many other disease areas, and a lot of study data, including the ones of the studies discussed here, are not publicly available. Partial, incomplete or delayed publication does not affect only industry sponsored studies, and although they have shared data with EMA, none of the cohort studies have granted independent access to their data to date.

Some general suggestions for future research in registry data include prospective data collection and embedding the mandatory post-marketing surveillance as registry based activity. Detailed data should be reported and overlap between different studies accounted for. Research protocols should be published before-hand for transparency and ideally harmonized through improved communication among different research groups. Planning and executing prospective individual patient meta-analyses has been shown to be a powerful tool in other disease areas (Cholesterol Treatment Trialists’ (CTT) Collaborators, 2005; Reade *et al*, 2010). With this in mind, data sharing agreements and an overarching mechanism to identify patients enrolled in multiple studies could be adopted. These collaborative and nationwide studies have increased patient numbers, but merging data from different studies with homogenous data will facilitate provision of more timely and more accurate answers to our questions.

Impact on assessment of immunogenicity of future concentrates

Five of the six studies we assessed show that in PUPs, the use of Kogenate/Helixate NextGen is associated with a variable increase in inhibitor rate compared to Advate, with the single exception being the SIPPET study. Three of the 6 studies show a variable increase in the inhibitor rate in recombinant versus plasma derived concentrates, RODIN

and EUHASS being the exceptions, and UKHCDO not providing evidence to the comparison. Some of the overall unadjusted differences in the estimates calculated across the 6 studies achieve statistical significance, which however does not rule out unmeasured confounding. For the first time in the history of haemophilia, it has been suggested that different “safe” concentrates are associated with a potential differential immunogenicity and this was possible in PUPs but not in PTPs. If we believe the results, we assume that the sophisticated analyses performed in the above studies accounted for the variability among different patients and that all different risk factors have been balanced out in the analysis. Consistently over several studies, groups of about 150 to 200 PUPs have been sufficient to measure the immunogenicity of different products with such a precision that in 3 studies the difference between Advate and Kogenate/Helixate NextGen has been significant (Gouw *et al*, 2013b; Collins *et al*, 2014) or borderline significant (Calvez *et al*, 2014).

In previously treated patients (PTPs), the EUHASS study observed 0.13 (CI 0.05-0.27) inhibitors /100 treatment years in patients treated with Kogenate/Helixate NextGen compared with 0.11 (CI 0.03-0.25) inhibitors/100 treatment years in patients treated with Advate (rate ratio 1.18). This difference was not statistically significant (i.e. there is still room for the ratio going in the opposite direction), but the average (best) estimate is an 18% higher rate. The point we want to make here is that we have no evidence about the transferability of PUPs results to PTPs – we cannot say that they apply, but we cannot even say they don’t apply. In PTPs, where the rate of inhibitors development is about 30 times lower, many times more patients would be required to have the same precision, i.e. several thousand patients per group. These numbers are impossible to reach.

The SSC subcommittee of the ISTH (Dimichele *et al*, 2012) has recently updated the guidance for designing clinical trials for new factor concentrate registration. Beyond the different approach to calculate the required sample size, the recommendation reiterates the previous ISTH SSC recommendation (White *et al*, 2001) to use PTPs to assess the immunogenicity of new products, excluding that the new products do not have neo-antigens making them enormously immunogenic. In fact, the one study on which both recommendations are based is a single relatively small cohort of PTPs where the immunogenic effect of a neo-antigen of the factor VIII product was observed (Peerlink *et al*, 1993). However, this approach is completely underpowered to assess smaller, but still important, differences in immunogenicity between factor concentrates. This is why concentrates that surpassed the ISTH SSC proposed test (like Kogenate/Helixate NextGen), did show different inhibitor rates in large PUP studies where the inhibitor rate is much higher (Gouw *et al*, 2013b; Calvez *et al*, 2014; Collins *et al*, 2014). Similarly, all recombinant products were deemed to be immunologically safe when tested in PTPs, while SIPPET suggests that they are more immunogenic when used in PUPs. Therefore we should be very cautious in assuming safety in PUPs for new products based on evidence generated in PTPs.

A second point to consider is the most effective and meaningful approach to data analysis. Whilst we have ample evidence that inhibitor development in PUPs occurs within the first 50 EDs, the concept of exposure day may generate problems in the analysis for several reasons. Firstly, not all exposure days are the same: five exposure days in a row, or spread over two months cannot be the same for the immune system;

also, the frequency and distribution of exposure days over time is not random, but driven by disease related events and by treatment decisions; finally, the concept of exposure day can be difficult to compare when patients are treated with the new extended half-life concentrates. For all these reasons, using ED as the time scale in a Cox regression can lead to conclusions unrelated to the underlying biology. We have two proposals to overcome this issue: first, preferentially use a non-time dependent approach (logistic regression or CART analysis) in addition or instead of Cox regression. Second, limit the population used for comparison of immunogenicity to patients on prophylaxis, where the natural time scale and the ED should be more strongly correlated, even if differences in treatment intensity can still take place. A more feasible analytical solution may be to stratify the analysis (patients treated with early prophylaxis or primarily treated on demand) and present the results by stratum together with the overall ones. It would be interesting to assess the effects of such a secondary analysis on the SIPPET study results, where treatment intensity varied widely.

The appropriate use and ethical consequence of the data presented is to re-think how we assess immunogenicity, rather than focus the debate on Kogenate/Helixate NextGen or plasma derived vs recombinant FVIII only. Limiting ourselves to studying two concentrates, would distract from the big picture of assessment of side effect of concentrates and appropriate use of the analytical power of large cohorts of PUPs.

Novel research designs to study immunogenicity in PUPs

The most efficient way to predict the future is by studying the past. After Wight and Paisley (Wight & Paisley, 2003) raised the possibility that plasma derived concentrates were less immunogenic than recombinant products, a meta-analysis and meta-regression showed that this could have been due to confounding (Iorio *et al*, 2010). Different research groups approached the research question from different perspectives: the FrenchCoag group started a prospective data collection, comparing the plasma derivative Factane with Advate (Goudemand *et al*, 2015). The PedNET group designed RODIN, which unfortunately accrued a low number of patients treated with plasma derived FVIII. Finally the SIPPET study randomized over 250 PUPs to plasma derived or recombinant factor concentrates. The single recurrent theme of all these studies, independently from our confidence in the value of their results, is that they explored differential immunogenicity in PUPs of factor concentrates proven to be safe in PTPs according to the ISTH SSC proposed approach (Dimichele *et al*, 2012). The large participation in and discussion around these studies clearly point out that the haemophilia community is searching for something more than largely underpowered registration studies. Rather than performing PUP studies for separate products, a randomized controlled registry design would be more efficient (Lauer & D'Agostino, 2013). By using existing registry for the data collection, and randomizing patients within this registry, such a design would cost several times less than a standard RCT, be several times faster and be more applicable because it is conducted as routine clinical practice in a large proportion of the population. How would a randomized clinical registry trial be designed to answer the question about differential concentrate immunogenicity? Each participating haemophilia centre would indicate which concentrates they feel confident using (the number of concentrates is not limited and can be changed as long as it is pre-specified). Any new PUPs requiring

treatment would then be randomized to one of the concentrates selected by that centre, and followed up until inhibitor development or until 50 or 75 ED with a common standardized approach, e.g. RODIN (Fischer *et al*, 2014b), or a similar data collection. At pre-specified points in the data collection (e.g. assembly of 150 PUPs in a specific subgroup), carefully planned statistical analysis, adjusted for all the relevant covariates as in some of the studies we reviewed above (Calvez *et al*, 2014) is performed. The multivariable analysis would account for the inter-patient variability and the randomization would account for residual confounding by unknown confounders. Implementing a randomized registry trial design would be a large undertaking, but it would be feasible and highly efficient (Lauer & D'Agostino, 2013). A first important question that would need to be addressed in such a design is that of the immunogenicity of new engineered products, and mainly extended half-life products and mimetics, for which there is insufficient evidence to draw any conclusions. Last but not least, these studies all together indicate that we need to radically change our strategy to prevent inhibitor development, moving our research focus from selection of the type or regimen of factor VIII to a more radical approach including manipulation of the immune-system to facilitate selective tolerisation to factor VIII (Bessede *et al*, 2014; Matino *et al*, 2015; Gupta *et al*, 2015; FVIII-targeting specific regulatory T-cell therapy: A novel translational approach for tolerance in Hemophilia A patients; Hu *et al*, 2007; Liu *et al*, 2014). From this perspective, shifting to the use of mimetics completely devoid of FVIII related epitopes might indeed dramatically change the landscape.

Implication for clinical practice

Plasma derived factor concentrates have been shown to be associated with fewer inhibitors than recombinant in a RCT and prospective cohort. However, the residual rate of inhibitors is still too high to be considered acceptable. Five of six studies (Gouw *et al*, 2013a, 2013b; Calvez *et al*, 2014; Collins *et al*, 2014; Fischer *et al*, 2015; Peyvandi *et al*, 2016; Marcucci *et al*, 2015) show that use of Kogenate/Helixate NextGen is associated with an increase in inhibitor rate when compared with Advate in PUPs. There are weaker, but again consistent, figures pointing to higher immunogenicity for BDD factor VIII when compared to Advate (mostly due by an excess in low titer inhibitors, see table 1). Advate, in turn, has been found to lead/ was still associated with to the development of inhibitors of 42% in a small controlled series of selected/high risk PUPs despite initiation of early low intensity prophylaxis (Auerswald, 2014), and to give about 47% more inhibitors than Factane in the FranceCoag cohort (Goudemand *et al*, 2015). Again, the rate of inhibitors with the best option would still be too high.

Until new evidence, generated as we have suggested above, will be available to guide practice, what can we recommend to the clinician having to decide which concentrate to use in PUPs? We do not think there is an easy and universal answer. But there are three consideration to make: i) all of the existing concentrates, when tested following the ISTH SSC recommendation (Dimichele *et al*, 2012), were proved to be safe; as a consequence, ii) when the clinician considers using any of the newly introduced product, she/he should know that it could turn out to be as immunogenic as Kogenate/Helixate NextGen in the future, if properly studied in PUPs; iii) at the population level, it might be wiser not to

use Kogenate/Helixate NextGen, if the choice is between Kogenate/Helixate NextGen and Advate, and to use plasma derived concentrates, if more weight is given to a possible reduction in the risk of inhibitors over that of blood borne infections; however, whether the tolerance developed to plasma derived factor VIII would transfer to recombinant factor when switching after 50 ED or instead some more inhibitor would develop thus eliminating the initial advantage remains to be explored (Iorio *et al*, 2012; Hay *et al*, 2015). For all these reasons, when presenting the risk of developing inhibitor to the individual patient (or to his family), the message should be that all concentrates have been associated with a risk as high as 40%, and unfortunately predicting the individual inhibitor risk remains difficult. Indeed, the best way to proceed might indeed be to invite the patient or his parents to contribute to reducing the uncertainty and help future patients by consenting to the randomized registry trial describe above, ideally including random assignment to trials with new investigational products.

Acknowledgments

AI, KF and MM conceived the paper and critically reviewed the literature; AI drafted the manuscript; KF and MM reviewed and finalized the manuscript'

Disclosures

In the last 3 years AI's institution has received project based funding via research or service agreements with Bayer, Baxalta, Biogen Idec, Grifols, NovoNordisk, Pfizer and Octapharma. In the last 3 years MM has acted as consultant to CSL Behring, Grifols and NovoNordisk. He has received support from Bayer for travel and accommodation to attend an overseas meeting. He is the project leader of EUHASS which receives funding from Bayer, Biostest, BPL, CSL Behring, Grifols, Kedrion, LFB, NovoNordisk, Octapharma, Pfizer, Shire (Baxalta) and SOBI. In the last 3 years KF received speaker's fees from Bayer, Baxter, Biostest, CSL Behring, Biostest Octapharma, Pfizer, NovoNordisk; performed consultancy for Bayer, Baxter, Biogen, Freeline, NovoNordisk and Pfizer; and has received research support from Bayer, Wyeth/Pfizer, Baxter, and Novo Nordisk

References

- Auerswald, G. (2014) The EPIC Study: a lesson to learn. *Hemophilia*.
- Baker, S.G. & Lindeman, K.S. (1994) The paired availability design: a proposal for evaluating epidural analgesia during labor. *Statistics in medicine*, **13**, 2269–78.
- Baker, S.G., Lindeman, K.S. & Kramer, B.S. (2001) The paired availability design for historical controls. *BMC medical research methodology*, **1**, 9.
- Bassler, D., Briel, M., Montori, V.M., Lane, M., Glasziou, P., Zhou, Q., Heels-Ansdell, D., Walter, S.D., Guyatt, G.H., STOPIT-2 Study Group, Flynn, D.N., Elamin, M.B., Murad, M.H., Abu Elnour, N.O., Lampropulos, J.F., Sood, A., Mullan, R.J., Erwin, P.J., Bankhead, C.R., Perera, R., et al (2010) Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*, **303**, 1180–7.
- Van den Berg, H., Ljung, R. & and on behalf of PedNet study group (2015) Can a ‘center effect’ be the reason for the higher frequency inhibitors for a second generation recombinant factor VIII product? *Journal of Thrombosis and Haemostasis*, **13**, 69–AS193.
- van den Berg, H.M., Pipe, S. & Ljung, R. (2016) Plasma products do not solve the inhibitor problem. *Haemophilia*, **1**–2.
- Berntorp, E. & Iorio, A. (2015) Reflections on the FranceCoag report on inhibitory antibodies to factor VIII in patients with severe hemophilia A. *Blood*, **0**, 18–20.
- Bessede, A., Gargaro, M., Pallotta, M.T., Matino, D., Servillo, G., Brunacci, C., Bicciato, S., Mazza, E.M.C., Macchiarulo, A., Vacca, C., Iannitti, R., Tissi, L., Volpi, C., Belladonna, M.L., Orabona, C., Bianchi, R., Lanz, T. V., Platten, M., Della Fazia, M. a, Piobbico, D., et al (2014) Aryl hydrocarbon receptor control of a disease tolerance defence pathway. *Nature*, **511**, 184–190.
- Björkman, S., Blanchette, V.S., Fischer, K., Oh, M., Spotts, G., Schroth, P., Fritsch, S., Patrone, L., Ewenstein, B.M. & Collins, P.W. (2010) Comparative pharmacokinetics of plasma- and albumin-free recombinant factor VIII in children and adults: the influence of blood sampling schedule on observed age-related differences and implications for dose tailoring. *Journal of thrombosis and haemostasis : JTH*, **8**, 730–6.
- Calvez, T., Chambost, H., ClaeysSENS-Donadel, S., D'Oiron, R., Goulet, V., Guillet, B., Heritier, V., Milien, V., Rothschild, C., Roussel-Robert, V., Vinciguerra, C. & Goudemand, J. (2014) Recombinant factor VIII products and inhibitor development in previously untreated boys with severe hemophilia A. *Blood*, **124**, 3398–3408.
- Cholesterol Treatment Trialists' (CTT) Collaborators (2005) Efficacy and safety of cholesterol-lowering treatment: Prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *Lancet*, **366**, 1267–1278.

- Collins, P.W., Palmer, B.P., Chalmers, E. a, Hart, D.P., Liesner, R., Rangarajan, S., Talks, K., Williams, M. & Hay, C.R.M. (2014) Factor VIII brand and the incidence of factor VIII inhibitors in previously untreated UK children with severe hemophilia A, 2000-2011. *Blood*, **124**, 3389–3397.
- Colvin, B.T., Hay, C.R., Hill, F.G. & Preston, F.E. (1995) The incidence of factor VIII inhibitors in the United Kingdom, 1990-93. Inhibitor Working Party. United Kingdom Haemophilia Centre Directors Organization. *British journal of haematology*, **89**, 908–10.
- Committee on Strategies for Responsible Sharing of Clinical Trial Data Board on Health Sciences Policy Institute of Medicine (2014) Discussion Framework for Clinical Trial Data Sharing : Guiding Principles , Elements , and Activities.
- Darby, S.C., Ewart, D.W., Giangrande, P.L., Spooner, R.J. & Rizza, C.R. (1996) Importance of age at infection with HIV-1 for survival and development of AIDS in UK haemophilia population. UK Haemophilia Centre Directors' Organisation. *Lancet*, **347**, 1573–9.
- Darby, S.C., Kan, S.W., Spooner, R.J., Giangrande, P.L.F., Hill, F.G.H., Hay, C.R.M., Lee, C.A., Ludlam, C.A., Williams, M., Haemophilia, U.K. & Doctors, C. (2007) Mortality rates, life expectancy, and causes of death in people with hemophilia A or B in the United Kingdom who were not infected with HIV. *Blood*, **110**, 815–825.
- Darby, S.C., Keeling, D.M., Spooner, R.J.D., Wan Kan, S., Giangrande, P.L.F., Collins, P.W., Hill, F.G.H. & Hay, C.R.M. (2004) The incidence of factor VIII and factor IX inhibitors in the hemophilia population of the UK and their effect on subsequent mortality, 1977-99. *Journal of thrombosis and haemostasis : JTH*, **2**, 1047–54.
- Dimichele, D.M., Blanchette, V. & Berntorp, E. (2012) Clinical trial design in haemophilia. *Haemophilia : the official journal of the World Federation of Hemophilia*, **18 Suppl 4**, 18–23.
- Dolgin, E. (2014) Oral solutions. *Nature*, **515**, S167.
- Drazen, J.M. (2015) Sharing Individual Patient Data from Clinical Trials. *N Engl J Med*, **372**, 201.
- EMA/PRAC/332348/2016 (2016) Inhibitor development in previously untreated patients with severe haemophilia A treated with recombinant factor VIII products. *44*, 14–15.
- Fischer, K. & Blatny, J. (2016) Do the SIPPET study results apply to the patients I treat? *Haemophilia*.
- Fischer, K., Iorio, A., Hollingsworth, R. & Makris, M. (2016) FVIII inhibitor development according to concentrate: data from the EUHASS registry excluding overlap with other studies. *Haemophilia*, **22**, e36–e38.
- Fischer, K., Lassila, R., Peyvandi, F., Calizzani, G., Gatt, A., Lambert, T., Windyga, J., Iorio, A., Gilman, E. & Makris, M. (2015) Inhibitor development in haemophilia

according to concentrate. *Thrombosis and Haemostasis*, **113**, 968–975.

Fischer, K., Lewandowski, D., Marijke van den Berg, H. & Janssen, M.P. (2011) Validity of assessing inhibitor development in haemophilia PUPs using registry data: the EUHASS project. *Haemophilia*, **1**–6.

Fischer, K., Ljung, R., Platokouki, H., Liesner, R., Claeysens, S., Smink, E. & van den Berg, H.M. (2014a) Prospective observational cohort studies for studying rare diseases: the European PedNet Haemophilia Registry. *Haemophilia*, **20**, e280–e286.

Fischer, K., Ljung, R., Platokouki, H., Liesner, R., Claeysens, S., Smink, E. & van den Berg, H.M. (2014b) Prospective observational cohort studies for studying rare diseases: the European PedNet Haemophilia Registry. *Haemophilia : the official journal of the World Federation of Hemophilia*, **20**, e280-6.

FVIII-targeting specific regulatory T-cell therapy: A novel translational approach for tolerance in Hemophilia A patients.

Goudemand, J., Rothschild, C., D'Oiron, R., Demiguel, V., Dalibard, V., Micheau, M., Lutz, P., Gruel, Y., Vinciguerra, C., H. C., Calvez, T. & and on behalf of FranceCoag Network (2015) Plasma-derived factor VIII products and inhibitor development in previously untreated boys with severe hemophilia A: report of the FranceCoag Network. *Journal of Thrombosis and Haemostasis*, **13**, 148–Abstract OR143.

Gouw, S.C., van den Berg, H.M., Fischer, K., Auerswald, G., Carcao, M., Chalmers, E., Chambost, H., Kurnik, K., Liesner, R., Petrini, P., Platokouki, H., Altisent, C., Oldenburg, J., Nolan, B., Garrido, R.P., Mancuso, M.E., Rafowicz, A., Williams, M., Clausen, N., Middelburg, R. a, et al (2013a) Intensity of factor VIII treatment and inhibitor development in children with severe hemophilia A: the RODIN study. *Blood*, **121**, 4046–55.

Gouw, S.C., van der Bom, J.G., Ljung, R., Escuriola, C., Cid, A.R., Claeysens-Donadel, S., van Geet, C., Kenet, G., Mäkipernaa, A., Molinari, A.C., Muntean, W., Kobelt, R., Rivard, G., Santagostino, E., Thomas, A. & van den Berg, H.M. (2013b) Factor VIII Products and Inhibitor Development in Severe Hemophilia A. *New England Journal of Medicine*, **368**, 231–239.

Gupta, N., Culina, S., Meslier, Y., Dimitrov, J., Arnoult, C., Delignat, S., Gangadharan, B., Lecerf, M., Justesen, S., Gouilleux-Gruart, V., Salomon, B.L., Scott, D.W., Kaveri, S. V., Mallone, R. & Lacroix-Desmazes, S. (2015) Regulation of immune responses to protein therapeutics by transplacental induction of T cell tolerance. *Science Translational Medicine*, **7**, 275ra21–275ra21.

Guyatt, G.H., Briel, M., Glasziou, P., Bassler, D. & Montori, V.M. (2012) Problems of stopping trials early. *Bmj*, **344**, e3863–e3863.

Hay, C.R. (1998) Factor VIII inhibitors in mild and moderate-severity haemophilia A. *Haemophilia : the official journal of the World Federation of Hemophilia*, **4**, 558–63.

- Hay, C.R.M., Palmer, B., Chalmers, E., Liesner, R., Maclean, R., Rangarajan, S., Williams, M. & Collins, P.W. (2011) Incidence of factor VIII inhibitors throughout life in severe hemophilia A in the United Kingdom. *Blood*, **117**, 6367–6370.
- Hay, C.R.M., Palmer, B.P., Chalmers, E.A., Hart, D.P., Liesner, R., Rangarajan, S., Talks, K., Williams, M. & Collins, P.W. (2015) The incidence of factor VIII inhibitors in severe haemophilia A following a major switch from full-length to B-domain-deleted factor VIII: a prospective cohort comparison. *Haemophilia*, **21**, 219–226.
- Hougaard, P. (1995) Frailty models for survival data. *Lifetime Data Analysis*, **1**, 255–273.
- Hu, G., Guo, D., Key, N.S. & Conti-Fine, B.M. (2007) Cytokine production by CD4+ T cells specific for coagulation factor VIII in healthy subjects and haemophilia A patients. *Thrombosis and Haemostasis*, **97**, 788–94.
- Iorio, A. (2016) Research and policy implications of a recently published controlled study in previously untreated hemophilia patients at high risk of inhibitor development. *Haemophilia*, **accepted**.
- Iorio, A., Halimeh, S., Holzhauer, S., Goldenberg, N., Marchesini, E., Marcucci, M., Young, G., Bidlingmaier, C., Brandao, L.R., Ettingshausen, C.E., Gringeri, A., Kenet, G., Knöfler, R., Kreuz, W., Kurnik, K., Manner, D., Santagostino, E., Mannucci, P.M. & Nowak-Göttl, U. (2010) Rate of inhibitor development in previously untreated hemophilia A patients treated with plasma-derived or recombinant factor VIII concentrates: a systematic review. *Journal of thrombosis and haemostasis : JTH*, **8**, 1256–65.
- Iorio, A., Puccetti, P. & Makris, M. (2012) Clotting factor concentrate switching and inhibitor development in hemophilia A. *Blood*, **120**, 720–727.
- Keeling, D., Tait, C. & Makris, M. (2008) Guideline on the selection and use of therapeutic products to treat haemophilia and other hereditary bleeding disorders. A United Kingdom Haemophilia Center Doctors' Organisation (UKHCDO) guideline approved by the British Committee for Standards in Haemat. *Haemophilia*, **14**, 671–84.
- Lauer, Mi.S. & D'Agostino, R.B. (2013) The Randomized Registry Trial — The Next Disruptive Technology in Clinical Research? *New England Journal of Medicine*, **369**, 1579–1581.
- Liu, C.L., Ye, P., Lin, J., Djukovic, D. & Miao, C.H. (2014) Long-term tolerance to factor VIII is achieved by administration of interleukin-2/interleukin-2 monoclonal antibody complexes and low dosages of factor VIII. *Journal of thrombosis and haemostasis : JTH*, **12**, 921–31.
- Lo, B. (2015) Sharing Clinical Trial Data. Maximizing Benefits, Minimizing Risk. *JAMA*, **313**, 793–4.

- Makris, M., Calizzani, G., Fischer, K., Gilman, E., Hay, C., Lambert, T., Lassila, R., Ludlam, C. & Mannucci, P. (2011) EUHASS: The European Haemophilia Safety Surveillance System. *Thrombosis Research*, **127**, S22-25.
- Mannucci, P.M. & Tuddenham, E.G. (2001) The hemophilias--from royal genes to gene therapy. *The New England journal of medicine*, **344**, 1773-9.
- Marchesini, E., Prisco, D. & Iorio, A. (2011) Recombinant Factor VIIa Concentrate versus Plasma-derived Concentrates for the Treatment of Acute Bleeding Episodes in Persons with Haemophilia and Inhibitors. *European Oncology & Haematology*, **7**, 140-2.
- Marcucci, M., Mancuso, M.E., Santagostino, E., Kenet, G., Elalfy, M., Holzhauer, S., Bidlingmaier, C., Escuriola Ettingshausen, C., Iorio, A. & Nowak-Göttl, U. (2015) Type and intensity of FVIII exposure on inhibitor development in PUPs with haemophilia A. *Thrombosis and Haemostasis*, **113**, 958-967.
- MASAC MASAC Recommendation #243 on SIPPET (Survey of Inhibitors in Plasma-Product-Exposed Toddlers): Results and Recommendations for Treatment Products for Previously Untreated Patients with Hemophilia A. 2016.
- Matino, D., Gargaro, M., Santagostino, E., Di Minno, M.N.D., Castaman, G., Morfini, M., Rocino, A., Mancuso, M.E., Di Minno, G., Coppola, A., Talesa, V.N., Volpi, C., Vacca, C., Orabona, C., Iannitti, R., Mazzucconi, M.G., Santoro, C., Tosti, A., Chiappalupi, S., Sorci, G., et al (2015) IDO1 suppresses inhibitor development in hemophilia A treated with factor VIII. *Journal of Clinical Investigation*, **125**, 3766-3781.
- McGilchrist, C. & Aisbett, C. (1991) Regression with frailty in survival analysis. *Biometrics*.
- MRC, UKCRC & NIHR (2015) GOOD PRACTICE PRINCIPLES FOR SHARING INDIVIDUAL PARTICIPANT DATA FROM PUBLICLY FUNDED CLINICAL.
- Nijdam, A., Altisent, C., Carcao, M., Cid, A., Claeysens-Donadel, S., Kurnik, K., Ljung, R., Nolan, B., Petrini, P., Platokouki, H., Rafowicz, A., Thomas, A., Fischer, K. & Haemat, P. and C. study groups. (2015) leading before prophylaxis in severe hemophilia: paradigm shift over two decades. *Haematologica*, **100**, e84-6.
- Peerlink, K., Arnout, J., Gilles, G., Saint-Remy, J.-M.M., Vermeylen, J., Peerlinck, K., Arnout, J., Gilles, J.G., Saint-Remy, J.-M.M. & Vermeylen, J. (1993) A higher than expected incidence of factor VIII inhibitors in multitransfused haemophilia A patients treated with an intermediate purity pasteurized factor VIII concentrate. *Thrombosis and Haemostasis*, **69**, 115-8.
- Peyvandi, F., Mannucci, P.M., Garagiola, I., El-Beshlawy, A., Elalfy, M., Ramanan, V., Eshghi, P., Hanagavadi, S., Varadarajan, R., Karimi, M., Manglani, M. V., Ross, C., Young, G., Seth, T., Apte, S., Nayak, D.M., Santagostino, E., Mancuso, M.E., Sandoval Gonzalez, A.C., Mahlangu, J.N., et al (2016) A Randomized Trial of Factor VIII and Neutralizing Antibodies in Hemophilia A. *New England Journal of Medicine*, **374**, 2054-2064.

- Reade, M.C., Delaney, A., Bailey, M.J., Harrison, D.A., Yealy, D.M., Jones, P.G., Rowan, K.M., Bellomo, R. & Angus, D.C. (2010) Prospective meta-analysis using individual patient data in intensive care medicine. *Intensive Care Medicine*, **36**, 11–21.
- Sack, B.K., Herzog, R.W., Terhorst, C. & Markusic, D.M. (2014) Development of gene transfer for induction of antigen-specific tolerance. *Molecular Therapy — Methods & Clinical Development*, **1**, 14013.
- Scott, D.W. (2014a) Inhibitors - cellular aspects and novel approaches for tolerance. *Haemophilia*, **20**, 80–86.
- Scott, D.W. (2014b) Taking a leaf from the book of oral tolerance. *Blood*, **124**, 1548–1549.
- Wang, X., Su, J., Sherman, A., Rogers, G.L., Liao, G., Hoffman, B.E., Leong, K.W., Terhorst, C., Daniell, H. & Herzog, R.W. (2015) Plant-based oral tolerance to hemophilia therapy employs a complex immune regulatory response including LAP+CD4+ T cells. *Blood*, **125**, 2418–2427.
- White, G.C., Rosendaal, F., Aledort, L.M. & Lusher, J.M. (2001) Scientific and Standardization Committee Communication Definitions in Hemophilia Recommendation of the Scientific Subcommittee on Factor VIII and Factor IX of the Scientific and Standardization Committee of the International Society on Thrombosis and Haem. *Thrombosis and Haemostasis*, **85**, 560.
- Wight, J. & Paisley, S. (2003) The epidemiology of inhibitors in haemophilia A: a systematic review. *Haemophilia*, **9**, 418–35.

Table 1- Characteristics of large epidemiological studies assessing the effect of different factor VIII products on inhibitor development in previously untreated haemophilia A patients.

Years	RODIN 2000-2010	UKCHDO 2000-2011	FranceCoag 1991-2013	EUHASS 2008-2012	IPD-MA 1935-2010	SIPPET 2010-2014	Unique patients	Overall Rates % (95% CI)	I ²
Total patients	574	407	402	198	712	251	2544		
Total inhibitors, n (%)	177 (31)	118 (29)	134 (33)	42 (21)	183 (26)	76(30)	730	28.6 (25.6 31.8)	64.9*
Inhibitors => 5 BU, n (%)	116 (20)	60 (15)	75 (19)	29 (15)	148 (21)	50(20)	478	18.5 (16.5 20.8)	46.5
Total Plasma Derived	88	NA	99	20	531	125	863		
Total inhibitors, n (%)	29 (33)	NA	20 (20)	5 (25)	115 (22)	29(23)	198	23.7 (19.9 28.0)	31.6
Inhibitors => 5 BU, n (%)	21 (24)	NA	12 (12)	3 (15)	86 (16)	20(16)	142	16.7 (13.8 20.0)	17.5
Total recombinant	486	407	303	178	181	126	1681		
Total inhibitors, n (%)	148 (30)	118 (29)	114 (38)	37 (21)	68 (38)	47(37)	532	31.9 (27.3 36.8)	75.9*
Inhibitors => 5 BU, n (%)	92 (19)	60 (15)	63 (21)	26 (15)	62 (34)	30(24)	333	20.5 (15.8 26.3)	85.1*
Kogenate/Helixate	183	107	111	75	144	-	620		
Total inhibitors, n (%)	64 (35)	35 (33)	55 (50)	18 (24)	51 (35)	-	223	35.5 (28.7 43.0)	70.9*
Inhibitors => 5 BU, n (%)	40 (22)	19 (18)	28 (25)	14 (19)	48 (33)	-	149	23.6 (18.4 29.6)	62.5*
Advate	157	124	97	59	9	-	446		
Total inhibitors, n (%)	41(26)	29 (23)	33 (34)	13 (22)	3 (33)	-	119	26.9 (22.8 31.4)	4.7
Inhibitors => 5 BU, n (%)	25 (16)	14 (11)	20 (21)	7 (12)	3 (33)	-	69	15.9 (11.8 21.0)	35.0
Refacto AF/Xyntha	41	5	NA	19	NA	-	65		
Total inhibitors	15 (37)	3 (60)	NA	5 (26)	NA	-	23	35.6 (24.8 48.1)	0
Inhibitors => 5 BU	3 (7)	-	NA	4 (21)	NA	-	7	12.9 (4.3 32.7)	54.3*

Data from UKHCDO are after excluding overlap with RODIN; Data from FranceCoag are from data (Calvez *et al*, 2014) and (Goudemand *et al*, 2015) Goudemand 2016; Data from EUHASS are after excluding overlap with RODIN and FranceCoag; IPD-MA data are after excluding overlap with RODIN. Inhibitor rates were pooled using a random effect model. A mixed model analysis with

random effect for the study and fixed effect for the grouping factor was performed to explore difference among subgroups of patients, assuming independent variance for the different subgroups of different studies. The difference between plasma derived and recombinant product was significant for all inhibitors ($Q = 6.552$, $P=0.010$) but not for HR inhibitors ($Q = 1.593$, $P=0.207$); between Advate and Kogenate was significant for all inhibitors ($Q = 4.274$, $P=0.039$) and for HR inhibitors ($Q = 4.300$, $P=0.038$); between Advate and Xyntha was not significant for all inhibitors ($Q = 2.032$, $P=0.154$) or for HR inhibitors ($Q = 0.151$, $P=0.697$). All analyses were performed using Comprehensive Meta Analysis Ver 2.2.064

Table 2 – Critical appraisal of the study design and conduct of large epidemiological studies assessing the effect of different factor VIII products on inhibitor development in previously untreated haemophilia A patients.

Study, (Publ year)	Design*	Emrollment (years) Patients (Kogenate or Advate/Total)	Control Rate (#)	Risk difference (@)	Notes
RODIN (2013)	Retrospective, Prospective, Inception Cohort, Multinational	2000-2010 340/574	28.2	9	Increased risk with Kogenate found at post hoc analysis: hypothesis generation
UKHCDO (2014)	Retrospective, Inception Cohort, Single Country	2000-2010 300/407	23.8	11.3	Increased risk with Kogenate confirmed; increased risk with Refacto detected Evidence of temporal effect, and “centre” effect (higher rate in RODIN centres).
France C (2014)	Retrospective, Inception Cohort, Single Country	2000-2010 234/303	30	15	Increased risk with Kogenate confirmed “centre” effect, and “RODIN” effect (higher rate in RODIN centres).
EUHASS (2015)	Prospective, Registry Multinational	2009-2013 284/417	26.2	4.5	Increased risk with Kogenate NOT confirmed RODIN effect (higher rate in RODIN centres).
EAHAD IPD (2015)	Meta-analysis, Multinational	1994-2003 80/761	40	6.6	Increased risk with Kogenate NOT confirmed Direction of effect, Inconsistency
SIPPET (2016)	Randomized controlled trial, Multinational	2010-2014 NR/251	NR*	NR*	Increased risk with Kogenate NOT confirmed Centre effect

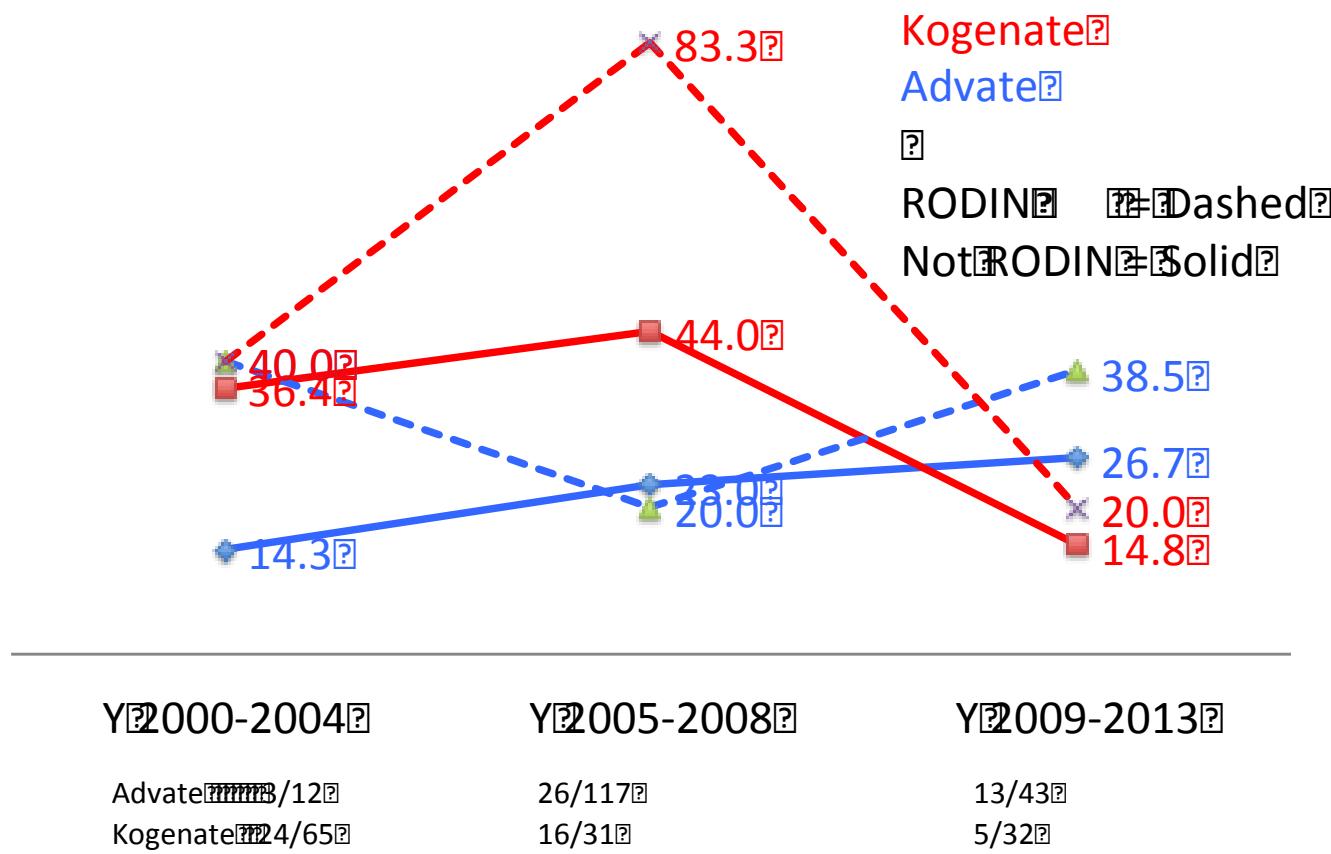
Legend to table 2.

= control rate, i.e event rate in the Advate group (%)

@ = risk difference, i.e absolute increase in risk in the Kogenate group (%)

We attributed IC as reported by the authors. However, some of the cohorts importantly deviated from being inception cohorts. The RODIN study is the closer to an inception cohort, having accounted for 88% of 648 eligible patients; the FranceCoag report excluded 437/741 patients because they started treatment before year 2000 or treated with plasma derived concentrates; the UKHCDO report described 86% of 468 patients, providing important details of the excluded patients as well; the EUHASS report accounts for 95% of centres and 95% of reported cases; the IPD meta-analysis is a pooled analysis of 4 centres, and does not report the proportion of patients accounted for; the SIPPET study reports on 82% of 303 eligible patients, but does not include a screening log (with some large centres having certainly seen many more PUPs than the one enrolled in the study), and the unusually high incidence of high risk mutation raises the suspect that a non-consecutive population was enrolled.

Figure 1. Trend over time of the rate of inhibitor development in previously untreated patients in the UKHCDO cohort



Legend to Figure 1. The figure displays the rate of inhibitor development in previously untreated haemophilia A patients in the UKHCDO cohort (ref). Rates are stratified depending on treatment (Advate or Kogenate), and depending on enrolment in the RODIN study. The rate increases for Advate and decreases for Kogenate over time, and the difference between Kogenate and Advate is much

larger in patients enrolled in the RODIN study. The difference in the rate of inhibitors for Kogenate treated patients between patients enrolled and not enrolled in the RODIN study is close to statistical significance ($p=0.08$). One possible interpretation for the large variability observed over time and across centers is selection bias.