Integrating clinicians' opinion in the Bayesian meta-analysis of observational studies: the case of risk factors for falls in community-dwelling older people

SILVIA DEANDREA[(1, 2)], EVA NEGRI[(1)], FABRIZIO RUGGERI[(3)]

ABSTRACT

BACKGROUND: despite the widespread application of Bayesian methods in meta-analysis, the incorporation of clinical informative priors based upon expert opinion is rare.

METHODS: a questionnaire to elicit beliefs about five risk factors for falls in older people was administered to a sample of geriatricians and general practitioners (GPs). The experts were asked to provide a point estimate and upper and lower limits of each relative risk. The elicited opinions were translated into different prior distributions and included in a Bayesian meta-analysis of prospective studies. Frequentist, Bayesian non-informative and fully Bayesian approaches were compared.

RESULTS: almost all the clinicians provided the requested information. In most cases, the variability across published studies was greater or similar to that across clinicians. Geriatricians provided more consistent estimates than GPs. When fewer studies were available, the use of the informative prior provided by geriatricians reduced the width of the credibility interval with respect to the frequentist or Bayesian non-informative approaches. Enthusiastic and skeptical priors led to results strongly driven by the prior distribution.

CONCLUSIONS: this study presents a feasible method for belief elicitation and Bayesian priors' assessment. The inclusion of external information showed to be useful when only few and/or heterogeneous studies were available from the literature.

Key words: Meta-analysis; Accidental falls; Bayesian analysis; Relative risk; Priors

(1) IRCCS "Mario Negri" Institute for Pharmacological Research, Milan, Italy
(2) Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy
(3) National Research Council of Italy - Institute of Applied Mathematics and Information Technology, Milan, Italy

CORRESPONDING AUTHOR: Silvia Deandrea, Department of Oncology, IRCCS "Mario Negri" Institute for Pharmacological Research, Via La Masa 19, 20156 Milan – Italy.
Tel: +39 02 39014-519. Fax: +39 02 33200-231.
e-mail: silvia.deandrea@marionegri.it

INTRODUCTION

Meta-analysis is a widely used method to synthesize evidence from multiple studies. The Bayesian approach provides a natural structure for many subtle issues that arise in meta-analyses, in particular for the use of available information through prior distributions [1] and the additional flexibility that derives from the adoption of Markov chain Monte Carlo methods.

However, prior distributions entertained in meta-analysis are generally non-informative and, even when informative priors are used, they frequently derive from the literature (evidence based priors) [2]. Conversely, the use of informative priors derived from expert opinion is more common in other study designs (i.e. clinical trials) [3-5] or in non-biomedical contexts (e.g. chemistry, veterinary, nuclear engineering) [5-7].

The scope of this work is to investigate the feasibility and the effect of enhancing observational findings using expert opinions from clinicians, through the development of an ad hoc questionnaire including a simple training section, and of a method to transform the elicited opinions into hyperparameters of prior distributions. Using the findings from a previously conducted meta-analysis on risk factors for falls [8] we compared the results using a) a frequentist approach, b) a non-informative Bayesian approach and c) an informative Bayesian approach with various priors.

METHODS

Bayesian approach to meta-analysis

A meta-analysis can be considered as a two level hierarchical model which relates the observed measure Y_k to the underlying effect in the kth study θ_k [9]. At the second level of the model, the θ_k s are related to the overall effect μ in the population from which all the studies are assumed to be sampled whereas σ^2 is the between-study variance. Y_k and θ_k are assumed to follow a normal distribution:

$$Y_k \sim N\left[\theta_k, s_k^2/n_k\right]$$
$$\theta_k \sim N\left[\mu, \sigma^2\right]$$

From a Bayesian perspective, s_k^2, μ and σ^2 are hyperparameters for which a prior distribution must be specified, as discussed later. For μ, a fully Bayesian approach was implemented, eliciting prior belief from experts.

Participants

A convenience sample of geriatricians and general practitioners (GPs) in Lombardy Region in Italy and Ticino Canton in Switzerland was contacted. Experts were recruited in the network of external collaborators of the Mario Negri Institute of Pharmacological Research. The search resulted in a list of 20 candidates, provided that not more than 2 were from the same group or institution. The following characteristics of the study participants were collected: age, sex, specialty, graduation year, year in practice treating older people, type of practice, history of statistical training, frequency of reading of fall risk factors papers, frequency of attendance to conferences about falls in older people.

Belief elicitation procedure

A questionnaire to elicit a relative risk (RR) was created, after a review of the methods used in the literature [10-11].

After the finalization of the tool, a standardized script was sent by e-mail to the participants (see Supplementary Materials 1). Participants were asked to specify, for a couple of community-dwelling older persons of the same sex and age but different with respect to the fall risk factor investigated, which one (i.e. affected or not by the risk factor) was more prone to fall in the subsequent year and the amount of risk increase. They were asked to indicate their response placing an 'X' on a line with risk increase ranging from 0.0 to 5.0. They were also told that 1.0 meant no difference, and values below 1.0 meant lower risk. Afterwards, they were asked to express the uncertainty around their estimate by placing an 'X' at the lower and upper limits of their belief.

The elicitation procedure for the first risk factor was supplied with instructions and examples, whilst for the subsequent risk factors the experts were asked to replicate the same procedure. At the end of the questionnaire, participants were told to compare the estimates and revise or correct them if needed. The questionnaire included five risk factors in the following order: benzodiazepines use, female sex, history of falls, urinary incontinence, antiepileptics use. For subjects aged 75 years and for more than 80 years old subjects the RRs were elicited separately.

Data from previous studies

The bases for our research were the papers included in a previously published

meta-analysis [8]. Briefly, from the results of a Medline and Embase search, we had selected original studies on risk factors for falls with the following inclusion criteria:

- At least 80% of the sample aged 65 years or older
- Prospective study design
- Sample size greater than 200 subjects
- At least 80% of subjects living in the community
- Number of subjects experiencing one or more falls during follow up as an outcome

Additionally, the reference lists of the previous reviews had been searched to identify studies that met the inclusion criteria. The outcome could be defined as:

- any faller - subjects who fell at least once during the follow up
- recurrent fallers - subject who fell at least twice during the follow up (within six months, or within one year, depending on the information given in the study)

From the 74 studies included in the previous meta-analysis, 31 articles [12-42] reporting odds ratios or RRs for the five selected risk factors and considering the outcome "any faller" were used in the current analysis. They were grouped according to the median or mean age of the sample (<80 vs. >80 years old); studies where information on mean or median age was not available were excluded.

Statistical analysis

Descriptive statistics was used to summarize the participants' characteristics and the elicited beliefs.

The experts were classified into 5 categories according to their characteristics:

- geriatricians
- general practitioners
- researchers - those who spend at least 10% of working time in research activity and/or reported to be trained in statistics during research activity
- informed - those who read fall risk factors papers and/or attend conferences about falls at least sometimes
- experienced - those who have at least 10 years working experience with older people.

Only the first two categories were mutually exclusive.

The frequentist pooled RRs were estimated using random effect models [43]. The statistical package used was RevMan, version 4.3.2 for Windows by the Cochrane Collaboration (http://ims.cochrane.org).

The Bayesian estimates of pooled RRs were obtained for risk factors for which more than 4 studies were available using different prior distributions for the hyperparameters. Our observations are log(RR)'s from the studies, whereas the posterior mean of the overall effect μ is used to obtain the log(RR) of the population. We assumed s^2_k known and replaced by the observed within-study variances. The between-study precision τ (equal to $1/\sigma^2$) was modeled using a vague prior testing three different gamma, i.e. (0.001, 0.001) (0.01, 0.01) (0.1, 0.1), or a uniform (0, 1) distribution, to reflect the high heterogeneity noticed in the classical random effects analysis. We used the vaguest prior for which convergence was obtained in a reasonable time, in order to balance computational efficiency and prior vagueness.

For μ we tested different priors:

- a vague non-informative prior, using a normal distribution centered at 0, with a large variance (10 000).
- two different clinical priors (i.e. one including all the elicited beliefs, the other including only the geriatricians' opinions), using a normal distribution with mean equal to the mean of the experts' point estimates, and a variance corresponding to the estimates' variance, following the method of moments proposed by Gaioni, Dey and Ruggeri [44].

The choice of a normal distribution, albeit very practical, was not taking in account the information provided on upper and lower limits and the consequent asymmetry; therefore we also considered a skew normal distribution [45].

The skew normal density function is:

$$f(x) = \frac{2}{\omega} \phi\left(\frac{x-\xi}{\omega}\right) \Phi\left(\alpha\left(\frac{x-\xi}{\omega}\right)\right)$$

where ξ, ω, and α are the location, scale and shape parameters respectively, φ is the normal density function and Φ is the normal cumulative distribution function.

- an enthusiastic prior, as the highest RR elicited from a single expert, using a skew normal distribution representing both the point estimate and the upper and lower limits
- a skeptical prior, as the lowest RR elicited from a single expert, also using a skew normal distribution.

WinBUGS version 1.4 (Medical Research Council Biostatistics Unit, Cambridge, England) (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml) and OpenBUGS version 3.1.2 (http://www.openbugs.net) were used to perform the Bayesian analyses with 2 separate chains with 20 000 Markov chain Monte Carlo iterations completed, after a 5 000 iteration 'burn in' period, for each pooled RR computed. We carried out graphical diagnostic for the convergence of the chains, as available in WinBUGS and OpenBUGS [46].

The reporting of the analysis and results is consistent with the ROBUST criteria [47].

RESULTS

Twenty physicians were asked to participate and 15 compiled the questionnaire. One subject gave an incomplete form, even after a telephone training session, and he was excluded from the analysis. The characteristics of the 14 experts included are reported in Table 1. There were 6 GPs and 8 geriatricians. The mean age of the sample was 44 years, and 50% of subjects included had more than 20 years experience after graduation. More than 60% of the sample had some statistical training. Almost all geriatricians read papers and attended conferences on falls at least sometimes, while GPs tended to be less informed about fall risk factors. We included 8 subjects in the category 'researcher', 10 in the category 'informed' and 9 in the category 'experienced'.

All the experts provided a point estimate and upper and lower limits for the five risk factors considered, separately for two age groups, except for one geriatrician who did not report estimates for antiepileptics in 75 years old subjects. For three risk factors

(benzodiazepines, incontinence, antiepileptics) for 80 years old subjects there were less than 4 studies available in the previous meta-analysis, so we excluded them from this study. The beliefs provided by the experts and the RR estimates from published studies for the most significant factors are shown in Figure 1 and in Supplementary Materials 2. In most cases, the variability between published studies was greater than (e.g. history of falls, 75 years old) or similar (e.g. history of falls, 80 years old) to the variability between experts. In both published studies and experts' beliefs some outliers were found (e.g. GP nr 2). The range of admissible values and the estimate provided by experts were sometimes very skewed (e.g. geriatrician nr 5 for incontinence). Confidence intervals (CIs) reported in published papers can be wider (e.g. history of falls, 75 and 80 years old) or narrower (e.g. incontinence) than the upper and lower limits reported by the experts. Figure 2 shows the distribution of RR estimates, according to experts' classification. For most risk factors, geriatricians provided more consistent estimates (e.g. female sex, 75 and 80 years old) with less outliers. Overall, the 'informed' subgroup did not provide estimates that were more consistent with published papers. The mean value of the central estimates was higher in informed, experienced and researchers than in geriatricians for almost all the risk factors considered.

The frequentist and Bayesian pooled RRs computed with non-informative and clinical priors are presented in Table 2 and Table 3. Enthusiastic and skeptical priors gave results that were too much driven by the prior distribution (data not shown). The application of a non-informative prior provided results that were very similar to the frequentist estimates (the gap is between -10% and +10% in all cases) with sometimes wider and sometimes narrower CIs; the CI for risk factors where the number of published papers was higher tended to be wider for the Bayesian estimate, like female sex. The application of an informative prior affected both the RR estimate and the CI. With the method of moments, results did not differ substantially from the frequentist ones. When we consider the prior provided by geriatricians only (with smaller variance), for risk factors with a small number of studies the CI was narrower: for example,

TABLE 1

CHARACTERISTICS OF THE 14 EXPERTS INCLUDED				
CHARACTERISTICS		**GPs (N=6)**	**GERIATRICIANS (N=8)**	**ALL (N=14)**
Male sex, n (%)	Male	4 (67)	4 (50)	8 (57)
Mean age (range)	-	52 (39-59)	40 (34-58)	44 (34-59)
Years from graduation, n (%)	More than 20 years	5 (83)	2 (25)	7 (50)
	10-20 years	1 (17)	4 (50)	5 (36)
	Less than 10 years	0 (0)	2 (25)	2 (14)
Specialty, n (%)	Geriatrician and/ or internist	2 (33)	7 (88)	9 (65)
	Other disciplines	2 (33)	0 (0)	2 (14)
	None	1 (17)	1 (12)	2 (14)
	Not reported	1 (17)	0 (0)	1 (7)
Type of practice*, n (%)	Public outpatients clinic	6 (100)	6 (75)	12 (86)
	Private outpatients clinic	2 (33)	1 (12)	3 (21)
	Hospital	0 (0)	4 (50)	4 (29)
	Long term care	0 (0)	3 (38)	3 (21)
	Nursing home	0 (0)	0 (0)	0 (0)
	Research	3 (50)	2 (25)	5 (36)
Experience in treating elderly people, n (%)	More than 20 years	5 (83)	1 (12)	6 (43)
	10-20 years	0 (0)	3 (38)	3 (21)
	Less than 10 years	1 (17)	4 (50)	5 (36)
History of statistical training*, n (%)	Post-graduate courses	3 (50)	1 (12)	4 (29)
	Personal study	3 (50)	3 (38)	6 (43)
	Research activity	5 (83)	2 (25)	7 (50)
	None	0 (0)	4 (50)	4 (29)
Frequency of reading of fall risk factors papers, n (%)	Very often	1 (17)	1 (12)	2 (14)
	Sometimes	1 (17)	6 (75)	7 (50)
	Rarely	2 (33)	1 (12)	3 (21)
	Never	2 (33)	0 (0)	2 (14)
Frequency of attendance to conferences about falls in older people, n (%)	Very often	0 (0)	2 (25)	2 (14)
	Sometimes	0 (0)	6 (75)	6 (43)
	Rarely	1 (17)	0 (0)	1 (7)
	Never	5 (83)	0 (0)	5 (36)

* Multiple answers
GP: general practitioner

for benzodiazepines (8 studies) the RR for the frequentist model was 1.36 (1.10-1.70) and the RR with geriatricians' prior was 1.34 (1.21-1.50), for antiepileptics (4 studies) the RR for the frequentist model was 1.88 (1.02-3.49) and the RR with geriatricians' prior was 1.94 (1.18-3.11). For other risk factors no appreciable change in the CI was noticed.

CLINICIANS' OPINION IN THE BAYESIAN META-ANALYSIS

FIGURE 1

RELATIVE RISK ESTIMATES FROM PUBLISHED STUDIES AND BELIEFS PROVIDED BY THE EXPERTS FOR HISTORY OF FALLS (75 AND 80 YEARS OLD), INCONTINENCE (75 YEARS OLD), BENZODIAZEPINES (75 YEARS OLD)
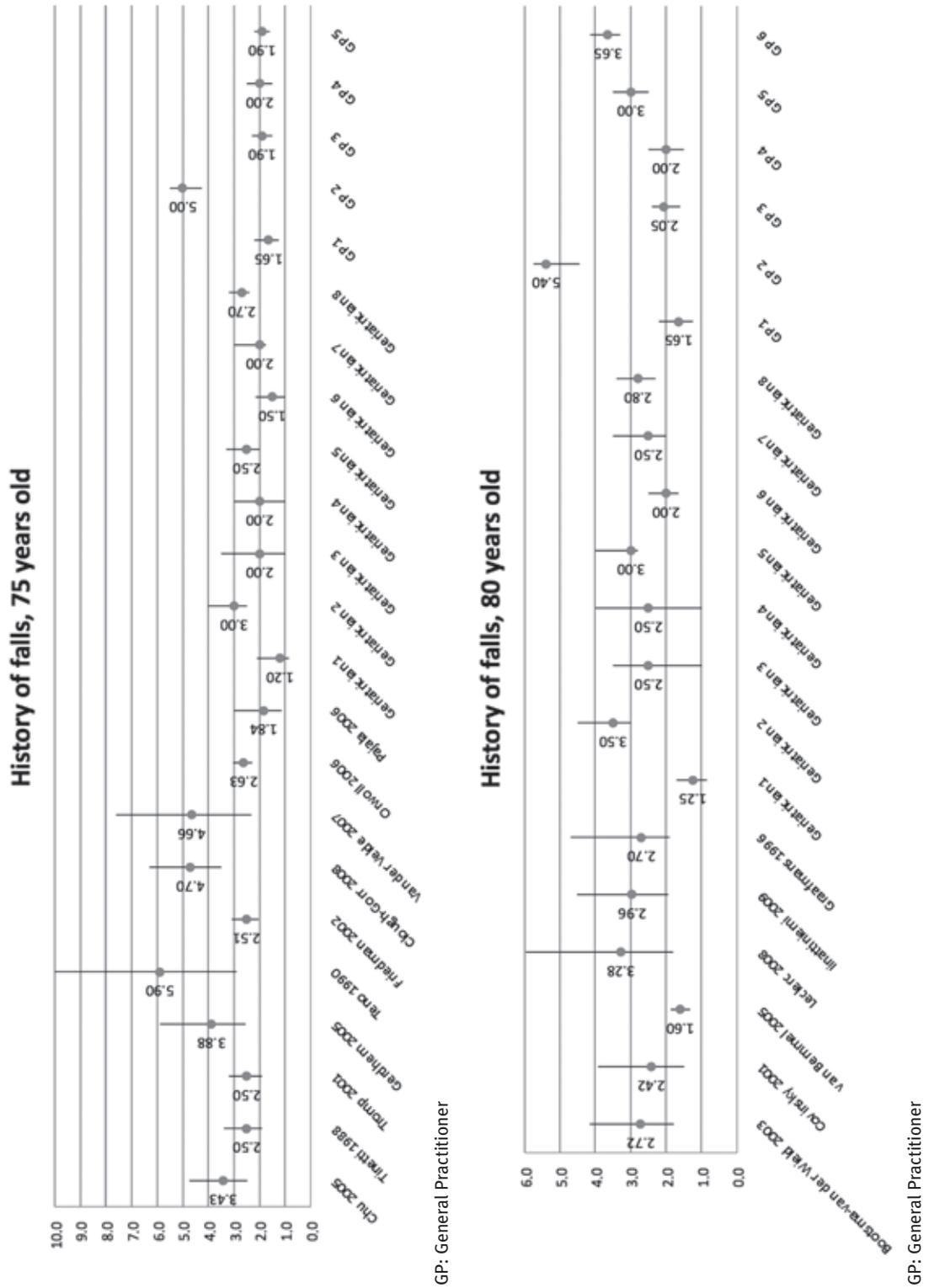

History of falls, 75 years old

GP: General Practitioner

History of falls, 80 years old

GP: General Practitioner

FIGURE 1 (CONTINUED)

RELATIVE RISK ESTIMATES FROM PUBLISHED STUDIES AND BELIEFS PROVIDED BY THE EXPERTS FOR HISTORY OF FALLS (75 AND 80 YEARS OLD), INCONTINENCE (75 YEARS OLD), BENZODIAZEPINES (75 YEARS OLD)

Incontinence, 75 years old

Benzodiazepines, 75 years old

GP: General Practitioner
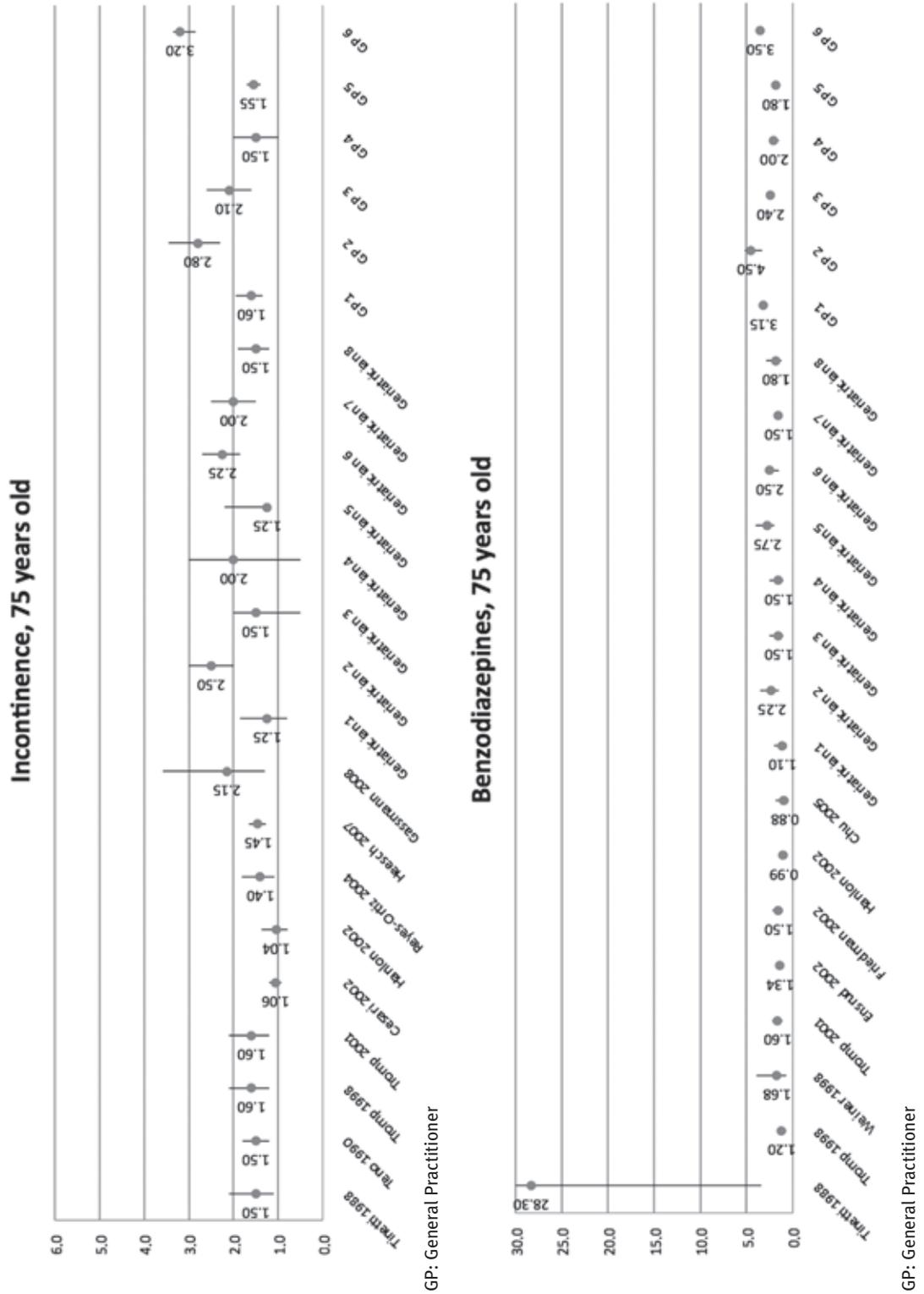
FIGURE 2

DISTRIBUTION OF RELATIVE RISK ESTIMATES, ACCORDING TO PUBLISHED STUDIES AND EXPERTS' CLASSIFICATION, FOR FEMALE SEX (75 AND 80 YEARS OLD), ANTIEPILEPTICS (75 YEARS OLD)
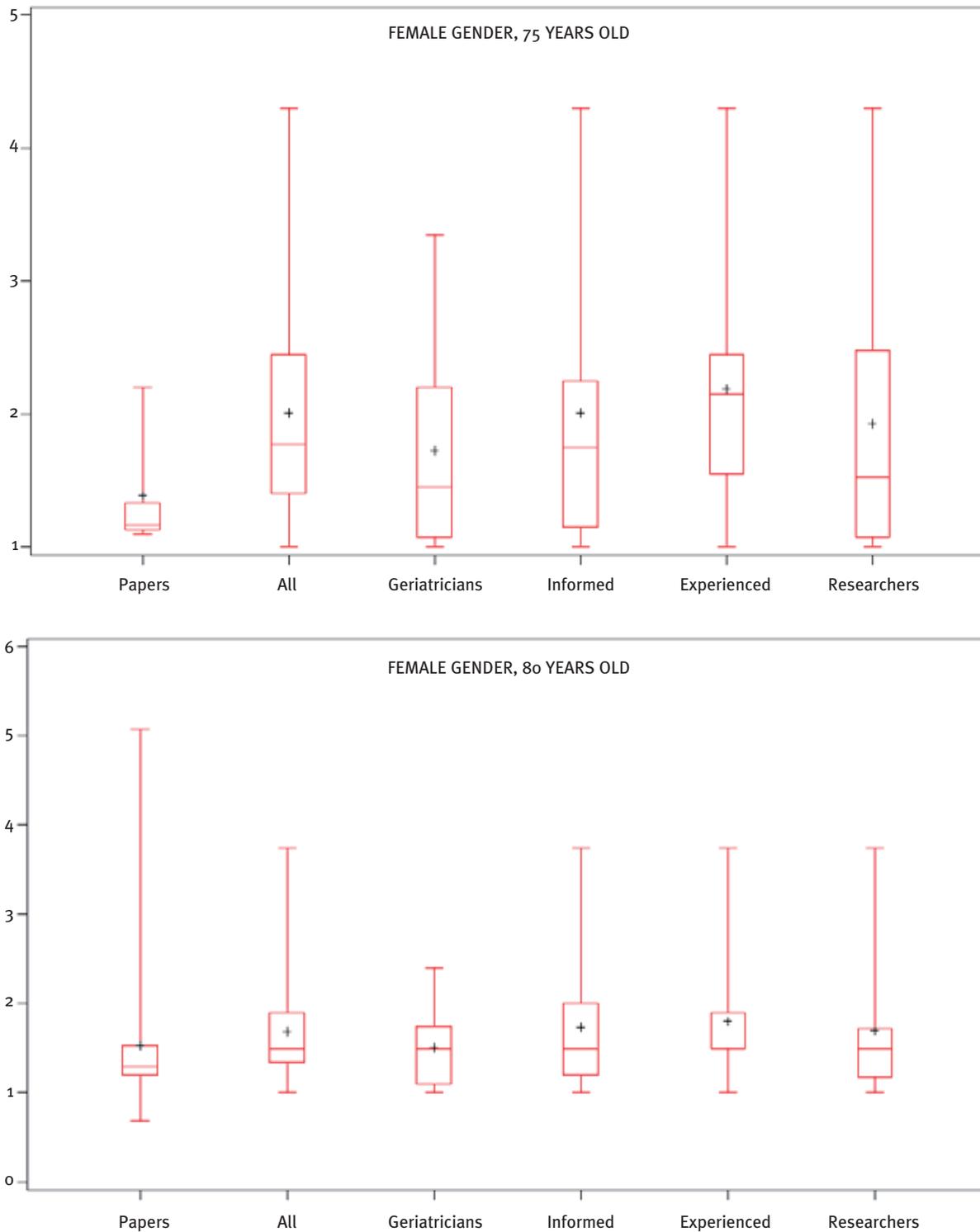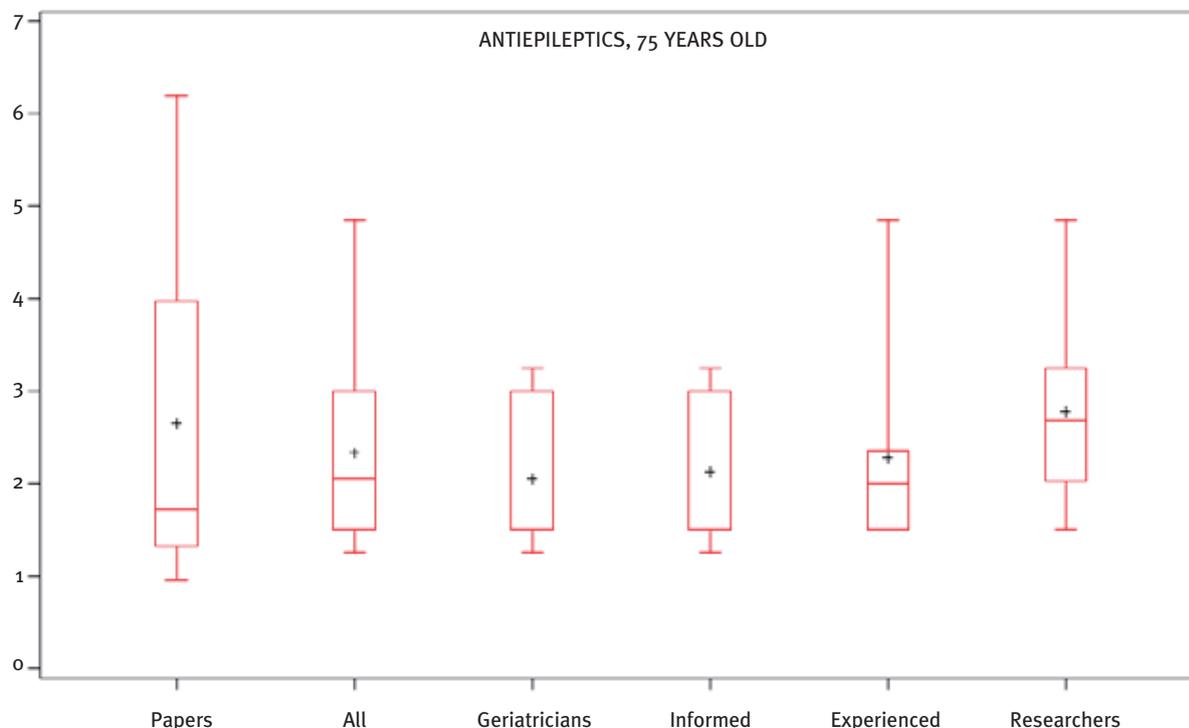

With the prior elicited from all the experts, in three cases there was a difference greater than ±20%. When we consider only the geriatricians' belief, the difference is smaller and in three cases we have a CI reduction >20%, that in two cases is accompanied by

FIGURE 2 (CONTINUED)

DISTRIBUTION OF RELATIVE RISK ESTIMATES, ACCORDING TO PUBLISHED STUDIES AND EXPERTS' CLASSIFICATION, FOR FEMALE SEX (75 AND 80 YEARS OLD), ANTIEPILEPTICS (75 YEARS OLD)

ANTIEPILEPTICS, 75 YEARS OLD

DISCUSSION

Our belief elicitation method presents an example of how expert opinion can be incorporated in research. We showed that this method of eliciting beliefs from clinicians is feasible and can be easily translated into a prior suitable for introduction in a random effects Bayesian meta-analysis.

This belief elicitation strategy included methods aimed at improving the validity and reliability of the elicited belief, as exemplified in the questionnaire proposed by Johnson et al. [11]. First, the questionnaire is easy to use with simple instructions, clear questions and response options, reducing the potential for invalid responses based on insufficient understanding of the task. The only physician who returned an inconsistent questionnaire differed from the other experts since he practiced in an isolated context. Second, we provided a guided elicitation procedure for

a coherence with frequentist estimates (e.g history of falls in 80 years old, antiepileptics).

the first risk factor, and training and provision of examples have been shown to improve reliability [48]. Third, we explicitly asked the participants to review their answers and verify whether these accurately reflected their beliefs, and to revise them at the end, if needed. The questionnaire was sent by e-mail with clear instructions and a telephone number to call if clarification was needed. Provision of feedback, verification, and opportunity for revision have also been shown to improve validity and reliability [7] although we did not evaluate them directly [11]. There is no general agreement about the number of participants required for a belief-elicitation study [3, 4, 11]. A systematic review [11] of elicitation methods found the median sample size of participants in belief-elicitation studies to be 11 (range: 1-298). Even if groups of experts are thought to perform better than a solitary expert [49, 50] it has been argued that experts are correlated sources of information because of their common education and exposure to similar literature [51], and even that the inclusion of an expert with beliefs identical to one already enrolled does not add to the range of beliefs collected

TABLE 2

	FREQUENTIST AND BAYESIAN UPDATED POOLED ODDS RATIOS (ORs)					
RISK FACTOR NUMBER OF PAPERS	**ODDS RATIO (95% CONFIDENCE INTERVAL /CREDIBILITY INTERVAL)**					
	FREQUENTIST RANDOM EFFECTS MODEL	**NON INFORMATIVE PRIOR**	**METHOD OF MOMENTS**		**SKEW NORMAL**	
			CLINICAL PRIOR: ALL	**CLINICAL PRIOR: GERIATRICIANS ONLY**	**CLINICAL PRIOR: ALL**	**CLINICAL PRIOR: GERIATRICIANS ONLY**
Female sex (75 years) n=17	1.32 (1.19-1.47)	1.40 (1.16-1.68)	1.41 (1.18-1.68)	1.41 (1.18-1.67)	1.51 (1.35-1.71)	1.42 (1.27-1.60)
Female sex (80 years) n=5	1.18 (1.03-1.37)	1.17 (1.09-1.27)	1.28 (1.01-1.78)	1.21 (1.05-1.41)	1.73 (1.41-2.04)	1.52 (1.34-1.75)
History of falls (75 years) n=11	3.06 (2.57-3.64)	3.02 (2.53-3.77)	2.86 (2.32-3.52)	2.80 (2.26-3.42)	2.41 (2.10-2.74)	2.49 (2.08-2.88)
History of falls (80 years) n=6	2.45 (1.82-3.28)	2.42 (1.82-3.31)	2.39 (1.89-3.18)	2.40 (1.89-3.12)	2.50 (2.16-2.87)	2.38 (1.97-2.87)
Incontinence (75 years) n=10	1.37 (1.22-1.53)	1.35 (1.23-1.54)	1.37 (1.24-1.57)	1.36 (1.24-1.56)	1.53 (1.33-1.81)	1.45 (1.28-1.70)
Benzodiazepines (75 years) n=8	1.36 (1.10-1.70)	1.25 (1.18-1.41)	1.39 (1.13-1.88)	1.34 (1.21-1.50)	2.01 (1.75-2.33)	1.60 (1.34-1.97)
Antiepileptics (75 years) n=4	1.88 (1.02-3.49)	2.01 (1.00-3.55)	2.03 (1.21-3.29)	1.94 (1.18-3.11)	2.10 (1.80-2.38)	1.87 (1.51-2.23)

in the study [52]. Some evidence suggests that the inclusion of participants with greater degree of competence improves the validity and reliability of the elicited response [50]. As experts we chose geriatricians and GPs, because geriatricians have specific competence on problems of older people, whereas GPs may have developed special competence on this topic because they directly take care of community-dwelling older people. We found that geriatricians gave consistent estimates with few outliers; they were also less overconfident than GPs, as shown by the narrower range of admissible values provided by GPs. This finding did not depend from the information about falls obtained reading papers or attending conferences.

Sampling experts from the population is a critical aspect, which can be addressed by proper stratification techniques and a larger sample size, whereas the correct method of sampling experts is also of concern. Although this was a convenient and not a random sample, with a limited sample size, we tried to represent different clinical settings choosing experts representing a relatively broad range of opinions of the target population, avoiding the inclusion of more than two experts from a single group or clinical unit.

We decided to elicit a RR instead of frequency of falls in the two groups (i.e. exposed and not exposed) because this method allowed the experts to provide a measure of the estimate precision. This option was also followed by other authors [11, 53]. However, as stated by Johnson et al. [11], there remains some ambiguity regarding the meaning of the range of probability. We had asked participants

TABLE 3

DIFFERENCES IN THE ESTIMATES OF ODDS RATIOS (ORs) AND 95% CREDIBILITY INTERVALS												
	DIFFERENCE IN THE ESTIMATE OF ODDS RATIOS						DIFFERENCE IN THE ESTIMATE OF 95% CREDIBILITY INTERVALS					
PRIOR	NEGATIVE GAPS			POSITIVE GAPS			NEGATIVE GAPS			POSITIVE GAPS		
	>20 (%)	20-11 (%)	10-0 (%)	1-10 (%)	11-20 (%)	>20 (%)	>20 (%)	20-11 (%)	10-0 (%)	1-10 (%)	11-20 (%)	>20 (%)
NON INFORMATIVE	0	0	5	2	0	0	2	0	1	2	1	1
METHOD OF MOMENTS												
CLINIC - ALL	0	0	3	4	0	0	0	2	0	1	1	3
CLINIC - GERIATRICIANS	0	0	4	3	0	0	2	1	0	3	0	2
SKEW NORMAL												
CLINIC - ALL	1	0	0	1	3	2	3	0	1	0	0	3
CLINIC - GERIATRICIANS	0	1	2	2	1	1	3	0	0	1	2	1

to indicate boundaries for which there is a very little probability that the true estimate could be greater or lesser. We do not know if this was interpreted by participants as plausible estimates, extreme values, or 95% credible intervals. We used this information in the skew normal model to determine its parameters, using the means of the upper and lower limits as quantiles of order 0.005 and 0.995 respectively. This model gave a better description of the experts' belief because it utilized the whole information provided. The method of moments, conversely, focused on the consistency across experts, neglecting the individual expert's uncertainty. Johnson et al. [54] recently used experts' beliefs on a survival probability by means of a multivariate Bayesian model. To our knowledge, however, this is the very first method proposed for the translation of experts' opinion in a prior for a meta-analysis. Sutton and Abrams [9] in their discussion of the application of Bayesian methods in meta-analysis stated that "to-date, many of the applications of Bayesian methods in meta-analysis have been to mirror classical random effects models". The inclusion of experts' opinions in the prior definition is a further step towards the full application of a Bayesian view in the meta-analytic context. Even if it has been already done in other study settings, like for instance clinical trials, this is a completely new approach with regards to

meta-analyses of observational studies, where non-informative priors have been usually used.

We decided to elicit information from experts directly, using questionnaires, and not to derive it from the literature, because the level of detail we needed (i.e. the interval of values of the relative risk that experts considered plausible) is seldom available in published reviews not based on meta-analysis. Moreover, in terms of epidemiological studies, the published literature was already included in the likelihood.

The use of a non-informative prior, corresponding to lack of external information, had no notable impact on the RR estimates as compared to the frequentist approach, as expected. The introduction of a subjective informative prior resulted in a marked difference in the RR estimates only when the skew normal distribution was applied. The geriatricians' opinions were more consistent and resulted in a more precise prior distribution for the method of moments. Both methods resulted in credibility interval shrinkage when the number of available studies was low. For antiepileptics, for example, only 4 studies were available from the literature and the RRs were heterogeneous, from 0.95 [32] to 6.20 [18]. The geriatricians' belief, conversely, varied from 1.25 to 3.25. The frequentist pooled RR was 1.88 (1.02-3.49). When the lack of external information was expressed with a non-informative prior, the credibility interval included 1.00. The addition

of external information from geriatricians did not have a great impact on the RR point estimate (1.94 and 1.87 with Gaussian and skew normal respectively), but the credibility interval was shrunk (1.51-2.23, with the skew normal distribution). Also when a skeptical prior was used, the RR estimate was lower, but still significant: 1.30 (1.05-1.66) (data not shown). It can be objected that the experts had previous information from the same papers included in the meta-analysis: this would result in a double counting of the same data, biasing the results toward whatever the data show and creating overconfidence [55]. In the case of antiepileptics, the data from the literature were too conflicting to create a background shared opinion about the strength of the association, and the 'informed' experts (those reading papers or attending conferences on falls) were not closer to the frequentist pooled RR than the other subgroups. This example shows how this method could be applied in decision-making contexts, when a decision must be made in a brief time, even when evidence is scarce and contradictory (e.g. in health technology assessment).

One of our aims was to compare the results of our approach to the frequentist one. In order to be able to perform a meaningful frequentist meta-analysis, we selected risk factors for which more than 4 studies were available. However, our method is even more useful when relevant and valid external frequency data are unavailable [56]. In this case, every source of information should be taken into account and a Bayesian meta-analysis can provide a framework for information synthesis. However, even when considerable literature data are available, skepticism derived from clinical judgment can be a reasonable position that should be formally examined [57]. Frequentist and non-informative Bayesian approaches tend to be closer, but this closeness will harm inference when these observations are biased and better external information is available [57]. Our analysis is enhanced presenting results from different priors to reflect different opinions about the parameter and performing a sensitivity analysis by seeing how results vary as the prior is varied; in this process, data representations can help one judge the priors that are credible enough to examine [55].

In conclusion, this is, to our knowledge, the first example of the use of a clinical informative prior based on experts' opinions assessed through a questionnaire for a random effect meta-analysis of observational studies. Our study should be viewed as an exploratory analysis on the potential of this approach, also considering the inclusion of a small convenience sample of experts. Nonetheless our results show that the elicitation method is feasible, and the experts' opinion can be successfully translated into a probability distribution that can be introduced in a Bayesian random effects model. The addition of external information to a frequentist meta-analysis showed to be useful in case of few and/or heterogeneous studies from the literature.

References

[1] Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation. New York: Wiley; 2003

[2] Woolcott JC, Richardson KJ, Wiens MO, et al. Meta-analysis on the impact of 9 medication classes on falls in elderly persons. Arch Intern Med 2009; 169: 1952-60

[3] Chaloner K, Rhame FS. Quantifying and documenting prior beliefs in clinical trials. Stat Med 2001; 4: 581-600

[4] Carlin BP, Chaloner K, Church T, et al. Bayesian approaches for monitoring clinical trials with an application to toxoplasmic encephalitis prophylaxis. Statistician 1993; 42: 355-67

[5] White IR, Pocock SJ, Wang D. Eliciting and using expert opinions about influence of patient characteristics on treatment effects: a Bayesian analysis of the CHARM trials. Stat Med 2005; 24: 3805-21

[6] Van der Fels-Klerx IH, Goossens LH, Saatkamp HW, Horst SH. Elicitation of quantitative data from a heterogeneous expert panel: formal process and application in animal health. Risk Anal 2002; 22: 67-81

[7] O'Hagan A, Buck CE, Daneshkhah A, et al. Uncertain judgements. Eliciting experts' probabilities. Chichester, UK: John Wiley & Sons Ltd; 2006

[8] Deandrea S, Lucenteforte E, Bravi F, et al. Risk factors for falls in community-dwelling older people. A systematic review and meta-analysis. Epidemiology 2010; 21: 658-68

[9] Sutton AJ, Abrams KR, Jones DR, et al. Methods for meta-analysis in medical research. New York: Wiley; 2000

[10] Johnson SR, Tomlinson GA, Hawker GA, et al. Methods to elicit beliefs for Bayesian priors: a systematic review. J Clin Epidemiol 2010; 63: 355-69

[11] Johnson SR, Tomlinson GA, Hawker GA, et al. A valid and reliable belief elicitation method for Bayesian priors. J Clin Epidemiol 2010; 63: 370-83

[12] Chu LW, Chi I, Chiu AY. Incidence and predictors of falls in the Chinese elderly. Ann Acad Med Singapore 2005; 34: 60-72

[13] Vellas BJ, Wayne SJ, Garry PJ, Baumgartner RN. A two-year longitudinal study of falls in 482 community-dwelling elderly adults. J Gerontol A Biol Sci Med Sci 1998; 53: M264-74

[14] Cesari M, Landi F, Torre S, et al. Prevalence and risk factors for falls in an older community-dwelling population. J Gerontol A Biol Sci Med Sci 2002; 57: M722-6

[15] Covinsky KE, Kahana E, Kahana B, et al. History and mobility exam index to identify community-dwelling elderly persons at risk of falling. J Gerontol A Biol Sci Med Sci 2001; 56: M253-9

[16] O'Loughlin JL, Robitaille Y, Boivin JF, Suissa S. Incidence of and risk factors for falls and injurious falls among the community-dwelling elderly. Am J Epidemiol 1993; 137: 342-54

[17] Tinetti ME, Inouye SK, Gill TM, Doucette JT. Shared risk factors for falls, incontinence, and functional dependence. Unifying the approach to geriatric syndromes. Jama 1995; 273: 1348-53

[18] Tromp AM, Smit JH, Deeg DJ, et al. Predictors for falls and fractures in the Longitudinal Aging Study Amsterdam. J Bone Miner Res 1998; 13: 1932-9

[19] Brown JS, Vittinghoff E, Wyman JF, et al. Urinary incontinence: does it increase risk for falls and fractures? Study of Osteoporotic Fractures Research Group. J Am Geriatr Soc 2000; 48: 721-5

[20] Friedman SM, Munoz B, West SK, Rubin GS, Fried LP. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. J Am Geriatr Soc 2002; 50: 1329-35

[21] Ensrud KE, Blackwell T, Mangione CM, et al. Central nervous system-active medications and risk for falls in older women. J Am Geriatr Soc 2002; 50: 1629-37

[22] Orwoll E, Lambert LC, Marshall LM, et al. Endogenous testosterone levels, physical performance, and fall risk in older men. Arch Intern Med 2006; 166: 2124-31

[23] Pajala S, Era P, Koskenvuo M, et al. Genetic factors and susceptibility to falls in older women. J Am Geriatr Soc 2006; 54: 613-8

[24] Salva A, Bolibar I, Pera G, Arias C. Incidence and consequences of falls among elderly people living in the community. Med Clin (Barc) 2004; 122: 172-6

[25] Teno J, Kiel DP, Mor V. Multiple stumbles: a risk factor for falls in community-dwelling elderly. A prospective study. J Am Geriatr Soc. 1990; 38: 1321-5

[26] Tromp AM, Pluijm SM, Smit JH, et al. Fall-risk screening test: a prospective study on predictors for falls in community-dwelling elderly. J Clin Epidemiol 2001; 54: 837-44

[27] Bootsma-van der Wiel A, Gussekloo J, de Craen AJ, et al. Walking and talking as predictors of falls in the general population: the Leiden 85-Plus Study. J Am Geriatr Soc 2003; 51: 1466-71

[28] Mahoney J, Sager M, Dunham NC, Johnson J. Risk of falls after hospital discharge. J Am Geriatr Soc 1994; 42: 269-74

[29] Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. N Engl J Med 1988; 319: 1701-7

[30] Weiner DK, Hanlon JT, Studenski SA. Effects of central nervous system polypharmacy on falls liability in community-dwelling elderly. Gerontology 1998; 44: 217-21

[31] Gerdhem P, Ringsberg KA, Akesson K, Obrant KJ. Clinical history and biologic age predicted falls better than objective functional tests. J Clin Epidemiol 2005; 58: 226-32

[32] Hanlon JT, Landerman LR, Fillenbaum GG, Studenski S. Falls in African American and white community-dwelling elderly residents. J Gerontol A Biol Sci Med Sci 2002; 57: M473-8

[33] Reyes-Ortiz CA, Al Snih S, Loera J, et al. Risk factors for falling in older Mexican Americans. Ethn Dis 2004; 14: 417-22

[34] van Bemmel T, Vandenbroucke JP, Westendorp RG, Gussekloo J. In an observational study elderly patients had an increased risk of falling due to home hazards. J Clin Epidemiol 2005; 58: 63-7

[35] Heesch KC, Byles JE, Brown WJ. Prospective association between physical activity and falls in community-dwelling older women. J Epidemiol Community Health 2008; 62: 421-6

[36] van der Velde N, Stricker BH, Roelandt JR, et al. Can echocardiographic findings predict falls in older persons? PLoS ONE 2007; 2: e654

[37] Clough-Gorr KM, Erpen T, Gillmann G, et al. Preclinical disability as a risk factor for falls in community-dwelling older adults. J Gerontol A Biol Sci Med Sci 2008; 63: 314-20

[38] Gassmann KG, Rupprecht R, Freiberger E. Predictors for occasional and recurrent falls in community-dwelling older people. Z Gerontol Geriatr 2009; 42: 3-10

[39] Leclerc BS, Bégin C, Cadieux E, et al. Risk factors for falling among community-dwelling seniors using home-care services: An extended hazards model with time-dependent covariates and multiple events. Chronic Dis Can 2008; 28: 111-20

[40] Iinattiniemi S, Jokelainen J, Luukinen H. Falls risk among a very old home-dwelling population. Scand J Prim Health Care 2009; 27: 25-30

[41] Bischoff-Ferrari HA, Orav EJ, Dawson-Hughes B. Additive benefit of higher testosterone levels and vitamin D plus calcium supplementation in regard to fall risk reduction among older men and women. Osteoporos Int 2008; 19: 1307-14

[42] Graafmans WC, Ooms ME, Hofstee HM, et al. Falls in the elderly: a prospective study of risk factors and risk profiles. Am J Epidemiol 1996; 143: 129-36

[43] DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7: 177-88

[44] Gaioni E, Dey D, Ruggeri F. Bayesian modeling of flash floods using generalized extreme value distribution with prior elicitation. Chil J Stat 2010; 1: 75-90

[45] Azzalini A. The skew-normal distribution and related multivariate families. Scand J Stat 2005; 32: 159-88

[46] Gelman A, Rubin DB. A single series from the Gibbs sampler provides a false sense of security. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM, Eds. Bayesian Statistics 4. Oxford University Press: Oxford, 1992; 625-31

[47] Sung L, Hayden J, Greenberg ML, et al. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. J Clin Epidemiol 2005; 58: 261-8

[48] Streiner DL, Norman GR. Health measurement scales. A practical guide to their development and use. 3rd edition. New York: Oxford University Press; 2003

[49] O'Hagan A. Eliciting expert beliefs in substantial practical application. Statistician 1998; 47: 21-35

[50] Clemen RT, Wolmark N. Combining probability distributions from experts in risk analysis. Risk Anal 1999; 19: 187-203

[51] French S. Aggregating expert judgment. Revista de la real academia de ciencias exactas, fisicas y naturales 2011; 105: 181-206.

[52] Kadane JB. Progress toward a more ethical method for clinical trials. J Med Philos 1986; 11: 385-404

[53] Johnson NP, Fisher RA, Braunholtz DA, et al. Survey of Australasian clinicians' prior beliefs concerning lipiodol flushing as a treatment for infertility: a Bayesian study. Aust NZ J Obstet Gyn 2006; 4: 298-304

[54] Johnson SR, Granton JT, Tomlinson GA, et al. Effect of warfarin on survival in scleroderma-associated pulmonary arterial hypertension (SSc-PAH) and idiopathic PAH. Belief elicitation for Bayesian priors. J Rheumatol 2011; 38: 462-9

[55] Greenland S. Weaknesses of Bayesian model averaging for meta-analysis in the study of vitamin E and mortality. Clinical Trials 2009; 6: 42-6

[56] Greenland S. Bayesian perspectives for epidemiological research: I. Foundations and basic methods. Int J Epid 2006; 33: 765-75

[57] Higgins JPT, Spiegelhalter D. Being sceptical about meta-analyses: a Bayesian perspective on magnesium trial in myocardial infarction. Int J Epid 2002; 31: 96-104

*