A nephrology guide to reading and using systematic reviews of observational studies

Pietro Ravani1,2, Paul E. Ronksley1,2, Matthew T. James1,2 and Giovanni F. Strippoli3,4,5,6,7

1O’Brien Institute for Public Health, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada, 2Libin Cardiovascular Institute of Alberta, University of Calgary, Calgary, Alberta, Canada, 3Diaverum Medical Scientific Office, Lund, Sweden, 4Renal Division, Fondazione Mario Negri Sud, Santa Maria Imbaro, Italy, 5School of Public Health, University of Sydney, Sydney, Australia, 6Department of Emergency and Organ Transplantation, University of Bari, Bari, Italy and 7Diaverum Academy, Bari, Italy

Correspondence and offprint requests to: Pietro Ravani; E-mail: pravani@ucalgary.ca

ABSTRACT

Systematic reviews are an ideal way of summarizing evidence from primary studies. While systematic reviews of randomized trials are broadly used to summarize benefits and harms of interventions, systematic reviews of observational studies are useful to summarize data on prevalence of risk factors in a population, distribution of outcomes or associations of different risk factors with outcomes. Also, systematic reviews can be useful to clarify potential reasons for conflicting data found in primary studies and explore sources of heterogeneity (variation in primary study data) to better understand epidemiological data and generate hypotheses for candidate interventions to improve outcomes. Summarizing data from observational studies in systematic reviews is a powerful tool to distil existing prognostic evidence in specific settings and inform patients and healthcare providers. In this article, we describe how to critically appraise the methods, interpret the findings and apply the findings of a systematic review of observational (prognostic) studies.

Keywords: CKD, epidemiology, evidence, meta-analysis, systematic review

INTRODUCTION

Mrs RP is a non-diabetic, African American 55-year-old woman, with Stage G2 chronic kidney disease (CKD; glomerular filtration rate between 60 and 90 mL/min/1.73 m²) secondary to hypertension, who comes to see you for routine care. Her proteinuria is 0.5–1 g/day and her blood pressure (130/80 mmHg) is well controlled with a thiazide and an angiotensin-converting enzyme inhibitor. She has a family history of cardiovascular disease and diabetes, but her personal history is negative for cardiovascular events. She has never smoked and her body mass index is 25 kg/m². You have discussed her cardiovascular risk, which is ~2% per year in people with CKD not treated with dialysis [1]. She is on a low sodium diet and follows instructions from the dietician in your clinic on how to maximize the intake of vegetables, fruits, whole grains and low-fat dairy products, and limit the intake of saturated and trans fats and sweets. While reviewing her lifestyle with you she asks you whether she is allowed to have one to two alcohol drinks per day considering her cardiovascular risk. She tried to find the answer on the Internet and found conflicting information so she needs your advice. You are not aware of any randomized controlled trial (RCT) to address the benefits and harms of alcoholic beverages consumption, and therefore, you are curious as to whether there are prognostic studies evaluating whether exposure to alcohol intake is associated with improved heart prognosis, and whether an RCT in this setting would be feasible.

In this review, we will explore Mrs RP question ‘Am I allowed to have one to two drinks per day considering my cardiovascular risk’. Because there are no RCTs testing the effect of alcohol consumption on cardiovascular outcomes, we will use available information from observational studies to discuss the meaning of association data and their implications. We will consider the following:

(i) Whether there is an association between alcohol consumption and the risk of cardiovascular events
(ii) The magnitude and certainty of this association if it exists?
(iii) How confident we are that such an association applies to this clinical scenario (i.e. a woman with risk factors for cardiovascular disease, including CKD and no previous history of cardiovascular events)?
(iv) What we should advise Mrs RP about her current alcohol consumption?
People coming to see us in clinic are often interested in discussing issues related to their lifestyle, in addition to objectives and expected benefits and harms of interventions. While RCTs are the only and ideal study design to address intervention questions, and provide unbiased estimates of treatment effects, evidence from RCTs may not be available or RCTs may in some occasion not be ethically or technically feasible. In these situations, authors may look at lower level of evidence (e.g. from cohort studies) keeping in mind that these show only association and not causation and may only be hypothesis generation, rather than used to wisely inform treatment choices. Alcohol consumption is an example of a cardiovascular risk factor largely debated in the medical literature and popular media. In the absence of RCTs testing benefits and harms of different doses of alcohol on cardiovascular outcomes, clinicians must interpret observational data when answering patients’ questions about taking alcohol and risk of cardiovascular disease. These would also drive useful RCTs in this setting. Given that patients and their family members have access to increasing scientific information on the Internet, and the literature in this area abounds, it is not surprising that Mrs RP might have found potentially conflicting data. Ideally a well-done systematic review of prognostic studies on this topic would be very helpful for our clinical discussion with the patient and to boost scientific research in the area. We are aware of a systematic review published in the British Medical Journal (BMJ) in 2011 [2] and wonder whether its results may help address the question of Mrs RP (Table 1). We will discuss how to interpret the methods and findings of a systematic review of prognostic studies in reference to this clinical case and decide how confident we are with the available data on the association between alcohol intake and cardiovascular events.

By the end of reading this article, you will be able to do the following:

(i) Assess whether a systematic review of prognostic studies applies to your clinical scenario;
(ii) Critically appraise the methods of a review;
(iii) Read a forest plot (the graphical summary of the associations from available studies);

| Table 1. Methodology and results of the systematic review by Dr Ronksley et al. [ ] |
|---------------------------------|-------------------------------------------------|
| **Systematic review PICO**      | General population; adults (≥ 18 years) without previous cardiovascular events |
| Participants/population         | Current alcohol consumers versus non-consumers (lifetime abstainers only as reference in studies excluding drinkers); number of drinks per day (12.5 g per drink) |
| Intervention and comparator     | Death from cardiovascular disease (any CVD event, including stroke) |
| Outcome                         | Incident ischemic or hemorrhagic stroke (fatal or non-fatal stroke) |
| **Systematic review methods**   | Death from coronary heart disease |
| Design                          | Death from stroke |
| Sources                         | Death from all causes (in included studies) |
| Search criteria                 | Prospective cohort studies (observational) |
| Search strategy                 | MEDLINE (1950 through September 2009) |
| Screening of titles and abstracts| EMBASE (1980 through September 2009) |
| Full-text assessment for inclusion| Plus bibliography of articles, cardiology conference proceedings and expert in the field (for unpublished studies) |
| Data extraction                 | No language restrictions |
| Data quality assessment         | Developed according to standard methods |
| Data synthesis                  | Performed by two authors (independently) |
| Stratified/sensitivity analysis  | Performed by two authors (independently) |
| Systematic review results       | Limited to study duration and confounding adjustment |
| Citations identified            | Random effects model. As opposed to fixed-effects models that make inference only about the studies included in the meta-analysis (How large is the true effect in the set of included studies), random-effects models provide an inference about a larger set of studies from which the selected studies included in the meta-analysis are assumed to be a random sample (How large is the average true effect in the larger population of studies) |
| Studies included                | According to study quality, participant characteristics, data reporting (measure of associations) |
| CVD mortality                   | Cumulative meta-analysis, small study effects (no diagnostics) |
| n = 4235                        | Incident coronary heart disease |
| n = 84 (including 34 only men; 6 only women; 44 both) | Stroke/mortality from stroke |
| Drunkers versus non-drinkers relative risk (RR) 0.75 (95% CI 0.7–0.8; high heterogeneity); RR ranging from 0.42 to 1.07 | In dose–response analysis, lowest risk associated with the consumption of 2.5–14.9 g/day (J-shape relationship) |
| In dose–response analysis, lowest risk associated with the consumption of 2.5–14.9 g/day (J-shape relationship) | Similar results |
| No association; however summary RR < 1 for ischemic stroke and RR > 1 for hemorrhagic stroke (RR 1.14; 95% CI 0.97–1.34) | Results unchanged since 1992 (for mortality from coronary heart disease) and 1999 (for mortality from cardiovascular disease) |
| RR = 0.87 (95% CIs 0.83–0.92)    | Death from all causes |
| Benefits > in women (few studies); results robust to number of adjustment variables in studies, follow-up duration, studies with lifelong abstainers | Cumulative meta-analysis |
| Results unchanged since 1992 (for mortality from coronary heart disease) and 1999 (for mortality from cardiovascular disease) | |
(iv) Interpret how associations may vary in different settings (subgroup and meta-regression analyses); and
(v) Apply the findings of a systematic review to address your clinical question.

**STEP 1: DOES THE SYSTEMATIC REVIEW EVALUATE OUR SPECIFIC CLINICAL QUESTION?**

In a previous article of this series, we learned that the first step for a successful search for evidence is to formulate the clinical question in a standard format known as the PICO (Population–Intervention/exposure–Comparison–Outcome) [3]. By matching our clinical question with the study question of the review we expect to find, or have found, we will maximize the chances to address our clinical question. In fact, when the authors of a systematic review design the protocol, they specify the PICO to identify, select, include and critically appraise all relevant studies in the review. In our case, a person with hypertension and mild CKD is the ‘population’, alcohol intake is the ‘intervention’ (called ‘exposure’ in observational studies), lower amount of alcohol or no alcohol intake is the ‘comparator’, and cardiovascular events are the ‘outcome’. Ronksley et al. compared alcohol drinkers (intervention/exposure) with non-drinkers (comparator) for the risk of overall mortality from cardiovascular disease, incident coronary heart disease, mortality from coronary heart disease, incident stroke and mortality from stroke and mortality from all causes when this was reported in the studies (outcomes). They included the results from longitudinal (prospective) cohort studies and restrict the question to studies of people without known cardiovascular disease (population). Although we may need to extrapolate the results reviewers observed in the general adult population (i.e. without CKD) to the case of Mrs RP who has CKD and hypertension, we can now consider this review relevant to our case.

**STEP 2: WAS THE SYSTEMATIC REVIEW CONDUCTED ACCORDING TO A PROTOCOL?**

As with any study, guidelines for good-quality reporting [4] recommend that systematic reviews be conducted following a protocol in which objectives and methods of the review are pre-specified. To enhance the quality and transparency of health research, authors of systematic reviews are now encouraged to develop the protocol of any review (i.e. reviews of clinical trials or observational studies) according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [5]. Older guidelines recommended authors of reviews of observational studies to follow the Meta-analysis of Observational Studies in Epidemiology (MOOSE) [6]. These protocols include definition of the primary outcomes of interest, description of data extraction, definition of outcomes and the methods for data analysis. Pre-specification of the review methods will reduce the risk of false-positive results (i.e. the probability that what was found it is not true and was detected by chance). The study by Ronksley et al. [2] was conducted according to a pre-specified protocol, designed according to recommended guidelines [6].

**STEP 3: WAS THE SEARCH FOR THE RELEVANT STUDIES EXPLICIT AND COMPREHENSIVE?**

While a systematic review of RCTs should be based on the large majority of available information (ideally all), this is not a strict requirement for systematic reviews of prognostic studies. In fact, because observational studies are inherently confounded (and therefore associations between exposure and outcome do not imply causation), precision of estimates of associations is not the main goal of meta-analyses of prognostic studies. According to some authors, however, the problem with the search of observational studies is largely related to the coding of this study type (as opposed to RCTs). Nevertheless, also these systematic reviews should report the way primary studies were searched to allow replication of findings, specify the sources of the data that the authors searched (e.g. MEDLINE or EMBASE) and the electronic strategy used to identify citations. Efforts to include conference proceedings or hand-searching are not as important as in systematic review of RCTs. In general, it is important that the evidence should be identified as thoroughly as possible and that individual studies should be commented upon, particularly when methodologic- al issues may be responsible for the differences in the key findings of the studies. Ronksley et al. [2] used a recommended strategy for systematic reviews of observational studies [7], based on three comprehensive search themes: identification of relevant terms related to the exposure of interest (Theme 1); identification of relevant outcomes (Theme 2) and identification of relevant study designs (Theme 3). To allow replication of the work, the authors reported details of the electronic search for each database.

**STEP 4: DID THE AUTHORS INCLUDE THE RELEVANT STUDIES IN THE REVIEW?**

Once the list of citations is available, reviewers identify and select the studies of interest according to the PICO. Because we are asking a question about prognosis, the most appropriate study design for inclusion is the observational design. To maximize the reliability of the review, the methods section of the study should confirm that the process for gathering all the relevant studies from the citations obtained in the search was robust, reliable and reproducible. In the review by Ronksley et al. [2], two authors independently reviewed all identified abstracts for eligibility and selected them for full-text review. The authors reported a high inter-rater agreement at the stage of abstract selection; disagreements were resolved by consensus. The same reviewers performed the full-text review of articles that met the inclusion criteria and articles with uncertain eligi- bility. Articles were retained if they met the inclusion criteria for study design (prospective cohort design), study population (adults ≥18 years old without pre-existing cardiovascular
disease), exposure (current alcohol use with a comparison group comprising non-drinkers), and outcome (overall cardiovascular disease mortality or atherothrombotic conditions, specifically incident coronary heart disease, coronary heart disease mortality, incident stroke or stroke mortality).

**STEP 5: DO ADDITIONAL RESTRICTIONS ON STUDY INCLUSION CHANGE THE VALIDITY AND APPLICATION OF THE RESULTS?**

Sometimes reviewers place restrictions in their search and selection, leading to exclusion of studies, for example, published in non-English language, studies of small size or short duration, or studies enrolling participants with some characteristics (men only or hypertensive only). As for the comprehensiveness of the search strategy also, this issue is more important in systematic reviews of RCTs than in systematic reviews of prognostic studies. While restrictions can reduce variability in the data (and heterogeneity), they can reduce generalizability. The review by Ronksley et al. [2] had broad enough inclusion criteria to increase the power of the review, generalizability of findings and the chance to investigate different types of events.

**STEP 6: WAS DATA EXTRACTION PERFORMED IN DUPLICATE AND ACCORDING TO A PRE-SPECIFIED METHODOLOGY?**

Once the studies have been selected for inclusion, data need to be extracted from each report to complete a qualitative assessment of the studies and a quantitative synthesis of the data. This process involves extraction of study characteristics (sample size, country, percentage male, mean age or age range, etc.), exposure and outcomes, and strictly adheres to the definitions specified in the study protocol. Sometimes extraction of information is challenging.

In the study by Ronksley et al. [2], the primary exposure variable was the presence of active alcohol drinking at baseline compared with a reference group of non-drinkers. The authors acknowledged that the non-drinker group might have included people who never drank and people who stopped drinking, inducing potential variability in the data. Therefore, they also extracted the data from studies that included only lifetime abstainers as the reference group and studies that distinguished former drinkers from non-drinkers. In addition, they extracted information on the amount of alcohol consumed, using grams of alcohol per day as the common unit of measure. Grams of alcohol per unit were derived when the number of drinks was reported and they standardized the portions of alcohol depending on whether beer, wine or distilled spirits was consumed. Finally, they categorized the volume of intake as <2.5 g/day (~<0.5 drink), 2.5–14.9 g/day (~0.5–1 drink), 15–29.9 g/day (~1–2.5 drinks), 30–60 g/day (~2.5–5 drinks) and >60 g/day (~>5 drinks).

The extraction of the outcomes was apparently less challenging as they included the presence or absence of death from cardiovascular disease (fatal cardiovascular or stroke events), incident coronary heart disease (fatal or non-fatal incident myocardial infarction, angina, ischaemic heart disease or coronary revascularization), death from coronary heart disease (fatal myocardial infarction or ischemic heart disease), incident stroke (ischemic or hemorrhagic events) or death from stroke. The authors extracted also the number of deaths from all causes from the included studies to conduct secondary analyses on the association between alcohol consumption and mortality from all causes.

**STEP 7: IS THERE A META-ANALYSIS? WHAT DO THE RESULTS MEAN?**

**Overall approach**

Once the studies have been selected and data extracted, reviewers can assess the evidence. Assessing the evidence involves summarizing available information ideally both qualitatively and quantitatively. While all studies included in a systematic review contribute to the qualitative assessment of the evidence (which we will discuss later), a variable number of studies can be included in a quantitative synthesis of the data, also known as meta-analysis. Sometimes this quantitative assessment cannot be done, because data are not amenable for pooling, e.g. because there is insufficient information or exposure and participant characteristics are different.

**Measures of associations**

As observed in other statistical methodologies, a number of different techniques are available for meta-analysis that will enable the ‘pooling’ of numerical data. These techniques include methods for obtaining pooled estimates of measures of association for binary outcomes (e.g. being alive or dead) including relative risks, odds ratios, incidence rate ratios or hazard ratios, and methods for continuous outcomes (e.g. levels of kidney function or parathyroid hormone), which produce pooled estimates of difference in means between exposure groups.

**Measures of uncertainty**

As all statistical analyses, meta-analysis procedures also report a measure of uncertainty around the ‘best estimate’ of the association between exposure and outcome, using 95% confidence intervals (CIs). A 95% CI provides a range of values within which the true measure of association (relative risk or difference in means) likely lies.

**Graphical representation**

These estimates from the individual studies can be represented in a graphical format called a forest plot, the graphical summary of a meta-analysis, named because of its tree-like appearance [8]. A forest plot displays the association between exposure and outcome from each study and an overall measure of association at the bottom. The forest plot displays the name of the study author and publication year, for example, the number of people with and without events for each study, and the values of the parameters from each study (e.g. relative risk with 95% CI). These parameters from studies
(the ‘whiskers’ in the plot) provide information about the precision of each estimate and of the overall summary measure of association, and about the consistency of the data (how similar the studies are). The size of the box around the point estimate obtained from the individual studies represents the weight it receives in the calculation of the overall pooled estimate. These two key results of a meta-analysis are summarized at the bottom of a forest plot, with a diamond centered on the pooled point estimate and points extended across the pooled 95% CI. Precision is summarized in the 95% CI around the summary point estimate (significant if the 95% CI does not cross the line of the null effect indicated by a vertical line in the plot). The forest plot reports also a measure of consistency in the data. Consistency is quantified in terms of heterogeneity (usually calculated as $I^2$) ranging from 0 to 100%. $I^2$ categories of 0–25, 25–50, 50–75 and 75–100% indicate heterogeneity that is potentially unimportant, moderate, substantial or considerable [9].

**STEP 8: ARE THE RESULTS FROM DIFFERENT STUDIES SIMILAR ENOUGH FOR THE SUMMARY ASSOCIATIONS TO BE MEANINGFUL?**

**Example of heterogeneity**

The extent of statistical heterogeneity between studies is important when considering the appropriateness and validity of producing a pooled estimate by meta-analysis. Consider the following example of two hypothetical meta-analyses of a prognostic factor that is associated with reduced risk of an adverse outcome (Figure 1). The figure reports the results of these two meta-analyses, which include six studies each. The two meta-analyses reach the same conclusion in terms of the ‘most likely value’ of the measure of association [relative risk (RR) = 0.75] and precision of the estimate (the 95% CI does not cross the line of null effect). Also, the total number of participants ($n = 8540$) and the total number of events ($n = 382$) are the same in the two meta-analyses. However, the two forest plots report different $I^2$ estimates: in one $I^2 = 0$; in the other $I^2 > 75%$.

**Meaning of heterogeneity**

The six studies of the left plot provide estimates of association between an exposure and outcome that vary within random chance; statisticians would say that these studies likely belong to the same population of studies. In this case, we can conclude that these studies are similar enough for us to believe the validity of the overall pooled estimate provided by the meta-analysis. This situation is more likely to occur in meta-analysis of RCTs, although it is not the rule. The six studies of the right plot provide estimates of the association between an exposure and outcome that vary beyond random chance; statisticians would say that they might not belong to the same population of studies. In this case we cannot conclude that the studies are similar enough for the overall pooled estimate provided by meta-analysis to be valid (i.e. one single ‘average’ of different measures may not be appropriate). This situation is almost the rule in meta-analysis of observational studies. There are several reasons potentially underlying the observed heterogeneity, including differences in study methods (bias) or differences in the study PICO (population, intervention/exposure, comparator or outcome). Important heterogeneity implies that the association between exposure and outcome may vary across patients or contexts.

**Exploring potential reasons for heterogeneity**

Stratified analyses and meta-regression are analytical techniques that allow the identification of factors (called ‘moderators’) that may explain some or most of the observed heterogeneity in the data (similar to the ‘effect modifiers’ in the analysis of variance). When studies vary because of the effects of moderators (e.g. if in men the RR is 0.7 and in women is 0.8, and the two estimates are statistically different, sex is a significant moderator), these analyses show that heterogeneity can be explained. The additional value in exploring sources of heterogeneity in meta-analyses of observational studies is related to the way summary estimates are calculated and the suboptimal control of bias in observational studies. Meta-analysis methods assign the greatest weight to studies that provide the most information based on the number of

![Figure 1](https://example.com/figure1.png)

**FIGURE 1:** Two samples of studies: in one meta-analysis estimates of associations are more similar to each other than in the other meta-analysis.
TREATMENT UNDERSTANDABILITY

In the last step.

This has implications for our discussion with Mrs RP, as we will discuss in the last step.

**Assessing bias**

In addition to providing a quantitative summary of the study data (meta-analysis), a systematic review is expected to provide an assessment and a summary of the quality of all included studies. Qualitative assessment involves consideration of the risk of different types of bias that may threaten the validity of the study findings. In observational studies, it is important to quantify the risk of each of the following types of bias: selection bias (i.e. extent to which the groups under comparison are derived from the same populations), participation bias (extent to which the study population represents the source and target populations), attrition bias (adequacy and completeness of follow-up), measurement bias (consistent exposure and outcome definition and classification across comparison groups), confounding bias (consideration of prognostic factors of interest in analysis) and analysis bias (appropriate analysis of the analysis and results reported). A variety of tools have been developed for summarizing the quality of observational studies [10]. In general, it is more useful to report these criteria for each study than to use summary scores [11], which do not provide transparent information about each domain.

**Importance of bias assessment**

These types of bias, as in RCTs, may result in overestimation or underestimation of the true association between exposure and outcome, and may have an important influence on our confidence in the prognostic implications of the systematic review. Unlike the case with RCTs, however, where there is empirical evidence that allocation concealment, blinding and inadequate reporting of losses to follow-up underlie different trial results, for these forms of bias there is a lack of empirical evidence that they underlie variability of results from different observational studies. Unfortunately, Ronksley et al. [2] assessed only one type of bias (confounding) in their review, and therefore, we do not know much about the quality of the included studies. Although the rating of the overall quality of evidence is not as important as it is in RCTs, again this limitation has implications for our discussion with Mrs RP, as we will discuss in the last step.

**STEP 9: HAVE THE AUTHORS APPRAISED THE RISKS OF BIAS IN THE INCLUDED STUDIES AND RATED THE EVIDENCE?**

Now we can use the data from the systematic review to address Mrs RP’s question about prognosis (‘Are 1–2 drinks per day associated with reduced cardiovascular risk?’). The systematic review provides several important pieces of information. First, we can start from the overall summary measure of association found in the review by Ronksley et al. [2]: the risk is ~25% lower in moderate alcohol consumers compared with non-drinkers (including lifetime abstainers), with the lowest risk in the category of people drinking 1–2 drinks per day. Because observational studies estimate associations and not effects (i.e. they do not prove causation), we prefer to speak in terms of ‘risk in people who drink alcohol’ as opposed to ‘effect of alcohol on risk’. This is because observational studies can neither fully account for measurable confounders nor consider unmeasurable factors determining alcohol consumption. We do not know whether the benefits observed in alcohol consumers are the consequence of alcohol consumption or other determinants of health associated with alcohol consumption. Only RCTs can address the question ‘does alcohol consumption reduce the cardiovascular risk’. Second, although Mrs RP belongs to the category of moderate alcohol consumers, and may be happy with the direction of the association found in the systematic review, we need to tell Mrs RP that the estimated association between alcohol consumption and lower risk is statistically significant (precise) but uncertain. The substantial heterogeneity found in meta-analysis of mortality from cardiovascular disease associated with alcohol consumption suggests that it is more important to consider the range of the estimates of the associated relative risk from individual studies (ranging from 0.42 to 1.07) rather than the pooled estimate. We do not know what explains this variation and which estimates of association best reflect the case of Mrs RP. As Ronksley et al. [2] did not find evidence of publication bias, we can assume that this range is not affected by lack of publication of small reports with more negative results. Third, we do not have enough information on the quality of the evidence. Using a similar approach to the grading system endorsed by the Cochrane collaboration for systematic review of RCTs (GRADE) [12], we could assess our confidence in the prognostic implications of the evidence based on the quality of studies. Fourth, the review found potential harms associated with alcohol consumption. Although the association with incident hemorrhagic stroke was imprecise (RR 1.14; 95% CI 0.97–1.34) and uncertain ($I^2 = 68$%; Dr Ronksley, personal communication), we need to let Mrs RP know that alcohol consumers may be at an increased risk for this event. Finally, Mrs RP might wish to know what these risk changes mean in terms of number of events in people with CKD. We can help Mrs RP weigh potential harms and benefits associated with alcohol consumption in the following way. Considering a baseline cardiovascular risk of 2% per year in people with CKD, we can tell her that 5 fewer people in 1000 experience a cardiovascular event among alcohol consumers than in non-alcohol consumers,
that is \((1 - 0.75)^*20\). Considering the risk of hemorrhagic stroke, which is \(\sim 0.2\%\) per year across all categories of CKD according to a recent study [13], we can tell her that \(\sim 3\) more people in 10 000 experience an intracranial hemorrhage among alcohol consumers than in non-alcohol consumers, i.e. \((1.14 - 1)^*20\).

**STEP 11: FINAL CONSIDERATIONS**

Considering the limitation discussed above, the evidence from this systematic review obviously cannot be directly integrated into clinical practice or public health messages without formal evaluation in pragmatic clinical trials. These trials are needed to properly address the questions of optimal patient selection, compliance, risks and benefits. In fact, the effects of alcohol consumption may vary across people’s characteristics, with some risk being higher among younger individuals (e.g. injury and violence) while others being potentially lower (e.g. mortality from cardiovascular disease) or uncertain (e.g. cancer) among older individuals [2].

**CONFLICT OF INTEREST STATEMENT**

None declared.

**REFERENCES**


Received for publication: 25.3.2015; Accepted in revised form: 19.5.2015