PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Lifetime Risks of Kidney Donation: A Medical Decision Analysis
AUTHORS	Kiberd, Bryce; Tennankore, Karthik

VERSION 1 - REVIEW

REVIEWER	Geir Mjøen Oslo University Hospital Norway
REVIEW RETURNED	06-Mar-2017

GENERAL COMMENTS	Thank you for allowing me to review this paper which is a relevant update of a paper previously published in 2013. In lack of adequate long term data on kidney donors, this paper is helpful in adding knowledge to those who have to assess risk in potential kidney donors. I have no comments regarding the choice of Statistical method and the display of the Markov modelling. However, the manuscript would be improved if the authors could add some more information. I would wish that the Authors described, at least summarily, the population cohort(s) that the model is based on. Limitations and representativity of these cohorts may affect results. Are there any specific limitations regarding these cohorts that the Reader should be aware of? How do the results of this paper affect recommendations for future research on outcomes among kidney donors? Is long follow-up time important for detecting adverse outcomes? If these results are valid for actual donor cohorts, how long time of observation would you recommend to those planning prospective studies on kidney donor survival? Please consider cumulative hazard curves instead of survival curves as this would make it easier to spot how the curves separate. Some of the e tables would benefit from a more elaborate description.

REVIEWER	Hassan Ibrahim University of Minnesota Minneapolis, MN 55455
REVIEW RETURNED	10-Mar-2017

GENERAL COMMENTS	This is certainly an important paper that attempts at better quantifying the risks associated with donation which is of huge benefit to all potential donors 1.Assumption number 3 is an issue as the only detailed account of why donors develop ESRD seems to indicate than an acute events

	was superimposed on the background of diabetes or hypertension(Kiddo et al) 2. The true proportion of donors with true measured GFR is only 14% 3. Having reduced GFR after donation is not the same as CKD. I do realize that they in one of the models assigned a lower risk to donors but I still struggle with that 4. Why would ESRD not be linked to reduction in lifespan? 5. How much of these risks are explained by being genetically related rather than due to donation? 6. Do these results apply only to 40 year old donors? This was not clear to me. 7. It would be important to actually state that this loss of life is trivial compared to the benefit the recipient incurs and that most donors would take this risk.

REVIEWER	Robert Steiner MD University of California at San Diego USA
REVIEW RETURNED	16-Mar-2017

GENERAL COMMENTS	This work uses data from many sources to estimate the life years lost associated with CKD and ESRD that might occur after living kidney donation. It addressed the mortality connected with uremia itself and does a service by calling attention to donor risk of advanced CKD. Relatively little is written about these risks. The study is a necessary attempt to predict events that are far into the future, even though most of our young candidates are still less that two decades post donation (Steiner AJT 2014). Given the difficulty of their task, results must be assessed against certain plausible epidemiologic and semiquantitative observations. The authors may wish to comment on the following concerns. Results are congruent with an emerging paradigm. Significant ESRD risks are delayed at least 25 years in this well-selected group, mirroring the slow, exponential accumulation of ESRD with age in the general population that causes the marked age-related variation in risks of donor candidates (their ref 11, Steiner AJT in press). 60 year olds have less risk than 30 year olds. Table 3 also confirms the newly described, substantial increase in ESRD risk from donation itself. Black risk is increased three fold as it is in the general population. Realistically assuming an inability to identify and deselect future diabetics increases ESRD risk threefold. Possibly because data from Grams' seminal NEJM study were used in the current analysis, similarities are expected. But certain problematic aspects of Grams might also remain. Grams required normal subjects to acquire a kidney disease and lose GFR quickly to reach ESRD during a short 6-7 year interval. Diseases that were present at entry were advantaged and caused secondary hypertension and proteinuria that "predicted" ESRD. Being normal at entry associated with safety, and Grams extrapolated this short-term

safety at a young age to very low lifetime risks. But a normal urinalysis and blood pressure do not predict the absence of new kidney diseases in later life; they do not predict a low risk of diabetes. In fact, the study was too short to allow classic development of any diabetic ESRD (Steiner NEJM 2016). Its short duration also resulted in only very low GFRs reaching ESRD and registering as risk factors. Higher GFRs all looked the same, and this may have caused a secondary analysis to estimate a relatively low GFR-related risk of donation that may have carried over to the current study. An 8-11 fold increase is estimated from more focused studies (their refs 5 and 6) and from population data (Steiner AJT 2014). Several studies suggest that we do not reduce post donation diabetic risk in donors (their refs 9 and 20). Adding a more realistic risk of diabetes in the current study is a valuable amendment to Grams risk estimates. Even so, given that ESRD is 40 years away for most young candidates, seeming to predict almost half of it at age 20 may reflect Grams' study design. The low risk predictions in ref 5 had similar problems (Steiner, Clin Trans 2016).

The general population faces high risks of mild to moderate lifetime reductions in GFR and progressively smaller risks of advanced CKD (their reference 14). If the unselected population lost 30 ml/min of GFR ("donated a kidney") its relatively high lifetime risk of mild to moderate CKD would become its risk of advanced CKD and ESRD (Steiner AJT 2014). If the greater two-kidney risks of a GFR of 30 ml/min become donor risks of ESRD, the donor selection examination has to predict those risks too.

A second implausibility is the prediction that only several months of life are lost to ill-fated candidates who develop ESRD. The model may have truncated the typical progression of CKD to ESRD. Most kidney diseases are very slowly progressing. In their reference 5, only 5% of lifetime ESRD developed in normal subjects by 8-9 years. Diabetes will take over a decade to destroy 30 ml/min of GFR, and other diseases will behave similarly (Iseki KI 2003). Loss of that same 30 ml/min at donation would result in the donor reaching ESRD at least 10 years sooner. These are the GFR-related risks that manifested quickly in ref 5. It is hard to believe that donors on dialysis for 10 years would live almost as long if they had not needed 10 years of dialysis. Parenthetically, the profession has believed in a substantial mortality from dialysis when preemptive transplantation is discussed.

More focus on diabetic risk would be welcome. It is singularly important for ESRD and mortality. A more realistic, adjusted diabetic risk for donor candidates is applied at age 40 in table 3. But 20 year olds will have much higher adjusted risks, as they are more likely to live to see diabetic ESRD. Diabetes takes about 25 years to reach ESRD. Much of the diabetes that begins after age 40 will not reach ESRD in a normal lifetime.

Their model correctly finds a reduction in risk for middle-aged candidates, but I am not sure how it addresses the following. In middle age, lifetime diabetes and kidney diseases are often present, progressing slowly, and can be identified to deselect ill-fated individuals. In the general population, this would seem to reduce the risk of a 55 year old well more than half (Steiner AJT 2014). But middle-aged risk is further reduced because requiring a minimum GFR of 80 ml/min for donation deselects the bottom third of the normal range of GFRs at age 55. The deselected candidates will

	have less renal reserve should a kidney disease begin and bear most of the residual epidemiologic risk, making the rest at substantially lower risk. We allow close to the entire normal range of GFRs in 20 year olds but would exclude the highest risk GFRs if they donated 35 years later. The current study, unlike Grams, does not maximize its well-conceived and individualized approach to emphasize individual risks. Young, black, and low normal GFR donors are at far higher ESRD and mortality risks. Even using the 3 fold increased adjusted diabetic risk in table 3, and their possibly low 3.5 fold risk of donation, 20-year-old nonblack donors have a 6.3% lifetime risk of ESRD. Blacks are at three times this risk. Even if lowering GFR by 30 ml/min (i.e., donation) causes only a 3.5 fold increase in risk, low normal GFRs (e.g., that are 30 cc less than high normal GFRs) will determine higher risks and high normal GFRs will determine reciprocally low individual risks, but group risk will stay constant and conceal this. One of the many strengths of this valuable effort is the capacity to individualize risk, and that should be prominent in the abstract, results, and discussion.

REVIEWER	Lois Kim Cambridge Centre for Health Services Research, Department of Public Health and Primary Care, University of Cambridge, UK
REVIEW RETURNED	25-May-2017

GENERAL COMMENTS	This work describes the updating of a previously published model of kidney donors, applying updated risks of mortality and other interim states. The study aims to provide model-based extrapolations to estimate sex and race-specific lifetime ESRD incidence and QALYs. Major comments 1. As with any disease progression model, the input parameters (along with model structure) are essentially what determine the outputs and results. It is therefore imperative that these are well-described both in terms of their source (and whether this is of sufficient quality, including any limitations regarding generalisability) and the form that they take. In this model, it is important to provide details about whether each transition probability differs by race/sex and whether they vary over time, for example, but this was not always clear. It would also be good to report the transition rates used, rather than just the relative increase in rates for donors as given in the appendix. Overall I did not feel there was sufficient detail to allow me to get a good feel for how the model results would be influenced by assumptions relating to the model inputs, which is key for interpreting the outputs. In looking at the key model results, I wanted to understand what was driving the higher loss of life-years in black donors for example, but this is difficult without knowing which transition parameters depend on race. I would also like to see more focus in the discussion from this perspective, looking at what aspects of the model are driving the between group differences. 2. There was also a lack of clarity about how the model was calibrated. It appears that the model was calibrated on the 15-year sex/race specific ESRD cumulative incidence. However, it was

	unclear how the various transition probabilities that feed into the ESRD state (e.g. the transition from diabetes to ESRD, the transition from CKD to ESRD) were estimated/calibrated to produce this. Furthermore, although the model may validate well on the 15-year incidence, there is no discussion of whether there was good validation over the whole 15 year period. 3. It is clearly a simplification to treat CKD as one combined state, but there is no justification or discussion relating to this modelling choice. CKD would be better modelled as the constituent stages (3a to 5), which would each have substantially different quality adjustment and transition rates to other states, including mortality. 4. It is not clear why Fig 1 does not include a state for hypertension and diabetes, as this is an extremely common combination of comorbidities. 5. The investigation of uncertainty does not appear to include any propagation of uncertainty in parameter estimates through to parameter outputs. For example, the estimates of the relative difference in transition rates or quality of life adjustments will be estimated with uncertainty; this combined (joint) uncertainty across all the input parameters can then be propagated probabilistically through the model in order to provide uncertainty intervals about the results. There are some confidence intervals presented in Table 1, but it is unclear how these have been derived. 6. Given that the mortality associated with CKD is generally not a direct result of the CKD itself (unless ESRD has been reached), it is an important assumption that the increased mortality for the CKD state is the same in the donor population as the non-donor population. This is acknowledged as a sensitivity analysis in terms of mortality, but it is not clear whether other aspects of this were also explored (progression to diabetes or proteinuria, for example). Related to this, the statements in the discussion (top of Pg 10) – "a significant percent of the loss of life was associated with CKD not ESRD" and "deaths during the CKD health state accounted for most of the projected increase in mortality and reduction in QALYs" need some further discussion. Minor comments Methods: You do not state the cycle length of your Markov model. Pg 10: "The mechanism by which low glomerular filtration rate CKD causes an increase in cardiovascular and all-cause mortality rate is not completely known" – is it known to be causal, or is it just a known association? Pg 11: "black individuals appear to transition from CKD to ESRD at a faster rate than white individuals"; isn't this dictated by the input parameters? Pg 12: the statement that "this would not be consistent with higher rates of ESRD in black males compared to white males in the general population" appears to contradict the estimates given in the previous sentence (that the added risk of ESRD was in 1 in 28 for white males compared to 1 in 22 for black males).

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
geir mjøen
Institution and Country, Oslo University Hospital, Norway
Thank you for allowing me to review this paper which is a relevant update of a paper previously published in 2013.
In lack of adequate long term data on kidney donors, this paper is helpful in adding knowledge to those who have to assess risk in potential kidney donors.
I have no comments regarding the choice of Statistical method and the display of the Markov modeling.
However, the manuscript would be improved if the authors could add some more information. I would wish that the Authors described, at least summarily, the population cohort(s) that the model is based on. Limitations and representativity of these cohorts may affect results. Are there any specific limitations regarding these cohorts that the Reader should be aware of?
Response: We agree. We expanded the limitations section. Population ESRD rates in the US are higher than European, therefore projected rates in kidney donors may also be higher.
How do the results of this paper affect recommendations for future research on outcomes among kidney donors? Is long follow-up time important for detecting adverse outcomes? If these results are valid for actual donor cohorts, how long time of observation would you recommend to those planning prospective studies on kidney donor survival?
Response: The fact that more life years are lost in health stats prior to ESRD suggests that follow up and intervention at the time of risk factor development (hypertension/DM) is important not only to prevent ESRD but also minimize life years lost. We added a sentence added in the discussion. Unfortunately studies to show impact of interventions in the donor population would need to be very large and very long as noted before (Kiberd, Transplantation Research 2013). Given evidence in the general population, periodic risk factor detection (hypertension, increase risk of ESRD) and intervention in donors as in the general population probably does not need additional study. At Continued long term observational follow up studies will provide better information to inform future live kidney donors.
Please consider cumulative hazard curves instead of survival curves as this would make it easier to spot how the curves separate.
Response: The hazard coefficient for progression was kept constant in donors compared to nondonors. The cumulative incidence ESRD curves appear non-proportional as death becomes a competing event earlier in donors and the development of diabetes mellitus (same in both) will have a greater relative impact on donors compared to non-donors. Ideally the cumulative survival figures should be in the main body of the paper but there was a limit.
Some of the e tables would benefit from a more elaborate description.
Response: We realize that some of the detail in the supplement makes no sense to people not familiar with modeling (Tables e6-e9). We did expand the description for these eTables.
Reviewer: 2
Reviewer Name: Hassan Ibrahim
Institution and Country: University of Minnesota, Minneapolis, MN 55455
This is certainly an important paper that attempts at better quantifying the risks associated with donation which is of huge benefit to all potential donors
1.Assumption number 3 is an issue as the only detailed account of why donors develop ESRD seems to indicate than an acute events was superimposed on the background of diabetes or hypertension(Kiddo et al)
Response: Assumption # 3. 'Donor and non-donors transition through a CKD state for at least one year before developing ESRD' . Theoretically individuals can develop AKI that can cause ESRD even from a normal state ie Goodpastures Disease/HUS/etc whether they donate or not. The model did not specifically model this option but rather for simplistic sense created the structure that most of those with a superimposed severe irreversible acute kidney injury would have been in a low GFR (<60

ml/min) for at least 1 year (non-donors and donors). These rates are low.
2. The true proportion of donors with true measured GFR is only 14%.
Response: We are not sure if this refers to mGFR pre-donation or post-donation. Almost all association population studies of CKD and mortality and progression to ESRD are done in patients using creatinine based formulas not measured GFR. Most donors are selected based on eGFR.
3. Having reduced GFR after donation is not the same as CKD. I do realize that they in one of the models assigned a lower risk to donors but I still struggle with that.
Response: I would agree that many of those with a low GFR are at a lower risk than patients in the general population. Our baseline risk of 1.3 is already lower that the general population of 1.7. When the risk was eliminated for isolated GFR the loss of life years was still important. Intuitively we believe that donors who develop ESRD have greater mortality than donors with normal kidney function. Along these same lines there must be a point at which a low GFR CKD is associated with higher mortality. The mechanism is unclear but the population studies have not shown that one gene diseases such as PCKD are free of the increase in death associated with CKD or ESRD. In your study (JASN 2016) of donors, an eGFR<60 ml/min was associated with a higher mortality risk (HR 4.6) as a time dependent co-variable independent of proteinuria, new onset hypertension and diabetes.
4. Why would ESRD not be linked to reduction in lifespan?
ESRD is linked to life span, the longer you live the more likely ESRD. See table 3 with non-biologic/biologic donors, female sex vs male sex etc. Higher rates of ESRD are associated with higher rates of death. The association is not perfect. Obesity, smoking, hypertension etc have higher relative risks for death than progression to ESRD.
5. How much of these risks are explained by being genetically related rather than due to donation?
Response: This is an important point but difficult to assess, since some non-biologic donors may have a family history of CKD/ESRD. We examined this in the table 3. The question is whether the disease that causes ESRD also increases mortality independent of kidney dysfunction. For example the increase in risk for mortality associated with a familial kidney disease such as Alports/Medullary cystic disease/etc may be different than a donor with increased familial risk of diabetes mellitus (donating to a parent with ESRD from diabetes mellitus). Diabetes mellitus increases both ESRD and all cause-mortality independent of kidney dysfunction. It is important to point out that rates of ESRD were not statistically different in the Muzaale JAMA 2014 analysis for related and unrelated donors even though the rates were >2fold higher in related donors. We used their overall rates in our paper. The cumulative incidence of ESRD was the essentially the same for the first 10 years (see figure 2c of Muzaale paper).
6. Do these results apply only to 40 year old donors? This was not clear to me.
Response: The data for donors was aggregate data with a median age of 40 from the reference Grams NEJM 2016. We do not know if the relative risks for developing ESRD over 15 years differ in younger or older donors. The observational studies show that younger donors <50 will have a lower 15 year cumulative incidence of ESRD with donation (Muzaale JAMA 2014) compared to donors >50, however cumulative incidence of ESRD is also likely to be lower in younger non-donors. We assumed that the relative increased risk with donation would be the same magnitude across all ages but this is currently an assumption.
7. It would be important to actually state that this loss of life is trivial compared to the benefit the recipient incurs and that most donors would take this risk.
Response: I would agree with you wholeheartedly, but that was not the intention or the perspective of this paper, these inferences can be made by others. We think the overall risk is small, but we want to highlight that it is not just ESRD that is important. We didn't have enough space in the main body to highlight the issue that donors might have higher discount rates and therefore even lower perceived long term risk.
Reviewer: 3
Robert Steiner MD
Institution and Country, University of California at San Diego, USA
This work uses data from many sources to estimate the life years lost associated with CKD and ESRD

that might occur after living kidney donation. It addressed the mortality connected with uremia itself and does a service by calling attention to donor risk of advanced CKD. Relatively little is written about these risks. The study is a necessary attempt to predict events that are far into the future, even though most of our young candidates are still less that two decades post donation (Steiner AJT 2014). Given the difficulty of their task, results must be assessed against certain plausible epidemiologic and semiquantitative observations. The authors may wish to comment on the following concerns.

Results are congruent with an emerging paradigm. Significant ESRD risks are delayed at least 25 years in this well-selected group, mirroring the slow, exponential accumulation of ESRD with age in the general population that causes the marked age-related variation in risks of donor candidates (their ref 11, Steiner AJT in press). 60 year olds have less risk than 30 year olds. Table 3 also confirms the newly described, substantial increase in ESRD risk from donation itself. Black risk is increased three fold as it is in the general population. Realistically assuming an inability to identify and deselect future diabetics increases ESRD risk threefold.

Response: Agree. There are prediction equations for the development of diabetes in the literature (Ibrahim JASN 2016). But we are not sure how useful these are in selecting eligible donors. Donation would have no effect on developing diabetes mellitus although there may be higher risks in those who are related and who donate up in age to someone with ESRD from diabetes mellitus.

Possibly because data from Grams' seminal NEJM study were used in the current analysis, similarities are expected. But certain problematic aspects of Grams might also remain. Grams required normal subjects to acquire a kidney disease and lose GFR quickly to reach ESRD during a short 6-7 year interval. Diseases that were present at entry were advantaged and caused secondary hypertension and proteinuria that "predicted" ESRD. Being normal at entry associated with safety, and Grams extrapolated this short-term safety at a young age to very low lifetime risks. But a normal urinalysis and blood pressure do not predict the absence of new kidney diseases in later life; they do not predict a low risk of diabetes. In fact, the study was too short to allow classic development of any diabetic ESRD (Steiner NEJM 2016). Its short duration also resulted in only very low GFRs reaching ESRD and registering as risk factors. Higher GFRs all looked the same, and this may have caused a secondary analysis to estimate a relatively low GFR-related risk of donation that may have carried over to the current study. An 8-11 fold increase is estimated from more focused studies (their refs 5 and 6) and from population data (Steiner AJT 2014). Several studies suggest that we do not reduce post donation diabetic risk in donors (their refs 9 and 20). Adding a more realistic risk of diabetes in the current study is a valuable amendment to Grams risk estimates. Even so, given that ESRD is 40 years away for most young candidates, seeming to predict almost half of it at age 20 may reflect Grams' study design. The low risk predictions in ref 5 had similar problems (Steiner, Clin Trans 2016).

Response: There is now follow up data to 25 years and these are showing accelerating rates of ESRD from Diabetes mellitus and hypertension in actual donors. We modeled, in the primary analysis, that DM, hypertension and proteinuria would develop at the same rate as in the general population. Lower rates were used in a sensitivity analysis. We agree the early ESRD is from new onset aggressive disease (idiopathic glomerulonephritis) and is likely unpredictable but seems to have a constant hazard over 25 years and will likely to be a constant contributor to ESRD in the long term. The impact of new onset diabetes mellitus causing ESRD in the model is still low at 20 years of follow up but becomes more prominent later.

The general population faces high risks of mild to moderate lifetime reductions in GFR and progressively smaller risks of advanced CKD (their reference 14). If the unselected population lost 30 ml/min of GFR ("donated a kidney") its relatively high lifetime risk of mild to moderate CKD would become its risk of advanced CKD and ESRD (Steiner AJT 2014). If the greater two-kidney risks of a GFR of 30 ml/min become donor risks of ESRD, the donor selection examination has to predict those risks too.

Response: We agree.

A second implausibility is the prediction that only several months of life are lost to ill-fated candidates who develop ESRD. The model may have truncated the typical progression of CKD to ESRD. Most kidney diseases are very slowly progressing. In their reference 5, only 5% of lifetime ESRD developed

in normal subjects by 8-9 years. Diabetes will take over a decade to destroy 30 ml/min of GFR, and other diseases will behave similarly (Iseki KI 2003). Loss of that same 30 ml/min at donation would result in the donor reaching ESRD at least 10 years sooner. These are the GFR-related risks that manifested quickly in ref 5. It is hard to believe that donors on dialysis for 10 years would live almost as long if they had not needed 10 years of dialysis. Parenthetically, the profession has believed in a substantial mortality from dialysis when preemptive transplantation is discussed.

Response: Models can only predict a population risk. Since most donors do not develop ESRD and most are later in life the average loss across the population is relatively small. You are correct that some will lose many years of life just as some people lose many years of life from cancer, suicide CVD etc and some lose very little from these. Those with cancer at age 40 will lose more life years than those at age 80. Most donors will not develop ESRD, those that do will lose life years and those who develop it in at younger ages will lose even more. Those donors that develop ESRD at younger ages will be transplanted with better overall survival. We added a sentence in the discussion to highlight your concerns.

More focus on diabetic risk would be welcome. It is singularly important for ESRD and mortality. A more realistic, adjusted diabetic risk for donor candidates is applied at age 40 in table 3. But 20 year olds will have much higher adjusted risks, as they are more likely to live to see diabetic ESRD. Diabetes takes about 25 years to reach ESRD. Much of the diabetes that begins after age 40 will not reach ESRD in a normal lifetime.

Response: Incidence rates for diabetes are increasing and this is a limitation. The model assumed similar rates as in the general population. See supplement for cumulative risks. As with ESRD/CKD the increase in DM rates occurs after age 40 so that there is not much difference in cumulative risk at age 20 versus age 40. The lifetime risk of diabetes mellitus (BMI 25-30) is 40% for males age 20 and only slightly lower at 37.9% for males age 40. (Lancet Diabetes Endocrin 2014, Gregg). We agree that those younger donors who develop diabetes before 40 will have greater risks of ESRD and even greater risks of ESRD with donation that those developing diabetes at later dates. This is incorporated in the model.

Their model correctly finds a reduction in risk for middle-aged candidates, but I am not sure how it addresses the following. In middle age, lifetime diabetes and kidney diseases are often present, progressing slowly, and can be identified to deselect ill-fated individuals. In the general population, this would seem to reduce the risk of a 55 year old well more than half (Steiner AJT 2014). But middle-aged risk is further reduced because requiring a minimum GFR of 80 ml/min for donation deselects the bottom third of the normal range of GFRs at age 55. The deselected candidates will have less renal reserve should a kidney disease begin and bear most of the residual epidemiologic risk, making the rest at substantially lower risk. We allow close to the entire normal range of GFRs in 20 year olds but would exclude the highest risk GFRs if they donated 35 years later.

Response: This is a real conundrum, but we can only model what we know in aggregate. Individual patient level data with long term follow up is required. Until this all we can do is provide best estimates on average donors.

The current study, unlike Grams, does not maximize its well-conceived and individualized approach to emphasize individual risks. Young, black, and low normal GFR donors are at far higher ESRD and mortality risks. Even using the 3 fold increased adjusted diabetic risk in table 3, and their possibly low 3.5 fold risk of donation, 20-year-old nonblack donors have a 6.3% lifetime risk of ESRD. Blacks are at three times this risk. Even if lowering GFR by 30 ml/min (i.e., donation) causes only a 3.5 fold increase in risk, low normal GFRs (e.g., that are 30 cc less than high normal GFRs) will determine higher risks and high normal GFRs will determine reciprocally low individual risks, but group risk will stay constant and conceal this. One of the many strengths of this valuable effort is the capacity to individualize risk, and that should be prominent in the abstract, results, and discussion.

Response: Although an individualized risk calculator is ideal we do not feel positions to offer this detail. A lot more data is required before we can use this clinically. We have concerns that the calculator by Grams shows inconsistencies especially with BMI. For instance we believe the impact of obesity on donation may be underestimated by the Grams calculator especially if obesity is

associated with down stream higher risks of diabetes mellitus. Added to discussion/limitation.
Reviewer: 4
Lois Kim
Institution and Country, Cambridge Centre for Health Services Research, Department of Public Health and Primary Care, University of Cambridge, UK
We appreciate this reviewer's expertise and insights. Making the model more sophisticated with incorporating more health states, uncertainty for all the estimates and better calibration to observed non-ESRD outcomes in donors over longer periods of follow up would be the ideal. However we believe this study includes sufficient health states to get a broad estimate of the potential loss of life associated with donation. We feel this study builds upon and adds value compared to the Gram's study for several reasons; including using loss of life years as an outcome rather than only ESRD, incorporating later events such as diabetes mellitus that were not fully accounted for in the Gram's study (See Dr Steiner's comments), potentially understanding the importance of follow up and interventions at risk factor development to reduce death prior to ESRD, and putting this risk in perspective with other diseases. It remains to be determined whether greater precision in the estimates are needed to inform donors. As noted above we do not feel that we have enough information to make an individualized calculator.
Major comments
As with any disease progression model, the input parameters (along with model structure) are essentially what determine the outputs and results. It is therefore imperative that these are well-described both in terms of their source (and whether this is of sufficient quality, including any limitations regarding generalisability) and the form that they take. In this model, it is important to provide details about whether each transition probability differs by race/sex and whether they vary over time, for example, but this was not always clear. It would also be good to report the transition rates used, rather than just the relative increase in rates for donors as given in the appendix.
Response: The inputs were imputed to match cumulative events. See figure in supplement for ESRD and CKD stage 3+ in the supplement. Incidence rates for diabetes mellitus were taken from the literature and refined to replicate lifetime cumulative events. Transition Rates were converted to probabilities. Tables were created at 5 -10 year intervals and the annual probabilities were interpolated by the Tree age program (see below response). We have attached the Tree age outputs in an Excel file for the primary analysis for the reviewer.
Overall I did not feel there was sufficient detail to allow me to get a good feel for how the model results would be influenced by assumptions relating to the model inputs, which is key for interpreting the outputs. In looking at the key model results, I wanted to understand what was driving the higher loss of life-years in black donors for example, but this is difficult without knowing which transition parameters depend on race. I would also like to see more focus in the discussion from this perspective, looking at what aspects of the model are driving the between group differences.
Answer: Blacks have higher baseline mortality (US vital statistics) than whites stratified by age. The cumulative stage 3A+ curves for blacks and whites are close (only slight shift to left for blacks), therefore the transition from normal to stage 3a+ (see reference by Grams AJKD 2013 Figure 2 and Table S1 for transition propabilities) are only slightly higher in blacks. The transition probability for example in our study at age 40 for an average general population black male was 0.001 per year but slightly less at 0.0009 per year for a white male for Normal to Stage 3. Our study included the possibility of arriving at CKD stage 3+ from a several prior states including normal, HTN, proteinuria, diabetes melltus etc, making a print out of transition probabilities very complex. The cumulative stage 4 and ESRD cumulative curves are shifted to the left much more for blacks relative to whites reflecting much higher rates of transition from stage 3a to ESRD. The transition probability from stage 3 to ESRD for example at age 40 for a general population black male was very high at 0.21 per year but only 0.08 per year for a white male. Transition rates change at each year to replicate the CKD and ESRD curves (see supplement). Separate tables for Normal GFR to CKD and for CKD to ESRD were created to account for differences among the 4 cohorts to replicate observed findings. There are also higher rates of diabetes mellitus especially in black females that are superimposed on these. Since

blacks are developing ESRD at higher rates and earlier, they have a higher loss of life from this state relative to whites. Since they transition from CKD to ESRD faster, they spend less time in CKD and are exposed to less cumulative mortality in this state. The observations that black males have higher rates of stage 4 and ESRD than white females but paradoxically lower prevalence of early stage CKD has been described by others (Muntner CJASN 2012 and Grams AJKD 2013).

2. There was also a lack of clarity about how the model was calibrated. It appears that the model was calibrated on the 15-year sex/race specific ESRD cumulative incidence. However, it was unclear how the various transition probabilities that feed into the ESRD state (e.g. the transition from diabetes to ESRD, the transition from CKD to ESRD) were estimated/calibrated to produce this. Furthermore, although the model may validate well on the 15-year incidence, there is no discussion of whether there was good validation over the whole 15 year period.

Response: The calibration was a multistage process. Both donors and non-donors were assumed to be free of diabetes, hypertension, CKD and proteinuria but could subsequently develop these risk factors/diseases at the same rate as in the general population (please see supplement). These rates were derived from population studies. For example the rate of developing diabetes mellitus in a 45 yo male was 0.008 per year but 0.016 for a 65 yo male (Gregg Lancet Diabetes Endocrin 2014). Slightly higher numbers were used for blacks to achieve the cumulative lifetime projected rates observed in subjects 20 years old and higher. A slow iterative process was required to ensure that cumulative risks of CKD, ESRD, and DM matched observed results. Previously used hypertension and proteinuria rates were used to achieve the known population prevalence. To create cumulative curves of ESRD, CKD and ESRD we assigned a transition reward in TreeAge that allowed us to sum cumulative rewards.

Once these model variable were set then a multiplier was empirically derived and applied to the transitions from normal to ESRD to match the 15 year ESRD cumulative incidence observed in non-donors and a higher value for the multiplier was empirically derived to match the 15 year cumulative ESRD in donors. The ESRD event incidence between 0 and 15 are extremely small. Observed event rates (see Kaplan Meier curves in Muzaale JAMA 2014) are extremely low and subject to considerable error. To examine this uncertainty we also changed the multiplier in donors to match 15 year ESRD the upper and lower bounds of the 95% confidence interval and show that the outputs change modestly (see Table 1).

3. It is clearly a simplification to treat CKD as one combined state, but there is no justification or discussion relating to this modeling choice. CKD would be better modeled as the constituent stages (3a to 5), which would each have substantially different quality adjustment and transition rates to other states, including mortality.

Response: We strongly agree that adding in Stages 3A, 3B, 4, and 5 (with and without renal replacement therapy) as well as stratified by each level of proteinuria would be more realistic, but greatly complicates the model by adding more health states with no ability to calibrate or validate. In addition the time spent in CKD stage health states >stage 3 is quite small (examine cumulative stage 4 and ESRD curves for blacks in Grams AJKD 2013). Adding health states would be possible if we had detailed individual patient level data in donors and non-donors with long term follow up. This would be the ‘ideal’ and a suitable project at a later date once the information became available. We added this limitation in the discussion.

4. It is not clear why Fig 1 does not include a state for hypertension and diabetes, as this is an extremely common combination of comorbidities.

Response: We agree and we assumed that patients with late onset diabetes would have hypertension. Although there may be some that do not, the risk was the same in both cohorts so the relative difference between the 2 would be small.

5. The investigation of uncertainty does not appear to include any propagation of uncertainty in parameter estimates through to parameter outputs. For example, the estimates of the relative difference in transition rates or quality of life adjustments will be estimated with uncertainty; this combined (joint) uncertainty across all the input parameters can then be propagated probabilistically through the model in order to provide uncertainty intervals about the results. There are some

confidence intervals presented in Table 1, but it is unclear how these have been derived.
Response: The uncertainty in the model was derived from Grams et al and is the 95% confidence interval for the cumulative ESRD event for their actual study. The multiplier that was empirically derived to match the 15 year cumulative ESRD rates (see response to 3#) was changed to match the upper and lower bounds of that confidence interval. We agree there is uncertainty in many of the estimates and Monte Carlo simulation could be used to establish confidence intervals. At this point we do not have access to reasonable 95% confidence intervals and the distribution for many of the variables in actual donors although they exist for the general population from which the donors are derived. Since the donors and the non-donors are the same people at baseline and we are examining differences, uncertainties for key variables will be in the same direction and of the same magnitude for each donor and nondonor pair. For instance if the added risk of diabetes mellitus is actually higher than imputed that risk will be underestimated for both donor and non-donor pair.
The 95% confidence interval for the point estimate for 15 cumulative ESRD rates were quite large, for example for black females was 0.85% (95% Confidence Interval, 0.37%, 1.35%). The other major uncertainty was the relative risk of mortality associated with CKD and the range of values was quite large (1.0 to 1.7). It would have been nice to use uncertainty estimates for CKD from the study by Grams (AJKD 2013) but these are not available. The uncertainty in the cumulative diabetes mellitus rates in the population was small (40 year old white male 39.9% [36.5%, 37.4%]Gregg 2014) and from an individual perspective is more likely to depend on other factors (BMI/family history). In our ideal patients the risk of diabetes mellitus was reduced to between 40% and 50% on the actual population incidence rates. Once more information is available then the model can be refined to include theses uncertainties. At present this is a first look at examining non-ESRD outcomes with the hope that with further study better individualized estimates can be generated.
We added in the supplement loss of life years in an otherwise 40 year old white male ideal donors considering the uncertainty in the point estimated for ESRD progression associated with obesity and smoking. We hope these one sensitivity analyses help the readers and reviewer understand the magnitude of the uncertainly in these estimates. In the case of obesity the uncertainty appears small, is larger for smokers. We also included an added caution in the limitations about the obesity. We assumed obese otherwise ideal non-donors would have low rates of diabetes, hypertension and proteinuria. This may not be the case. Short term follow up studies will not detect the impact of acquiring these risk factors and increasing the risk of ESRD within 15 years. It is likely that lifetime risks of ESRD will be higher in both donors and non-donors.
6. Given that the mortality associated with CKD is generally not a direct result of the CKD itself (unless ESRD has been reached), it is an important assumption that the increased mortality for the CKD state is the same in the donor population as the non-donor population. This is acknowledged as a sensitivity analysis in terms of mortality, but it is not clear whether other aspects of this were also explored (progression to diabetes or proteinuria, for example). Related to this, the statements in the discussion (top of Pg 10) – “a significant percent of the loss of life was associated with CKD not ESRD” and “deaths during the CKD health state accounted for most of the projected increase in mortality and reduction in QALYs” need some further discussion.
Response: We are not sure there is evidence to support the claim that mortality associated with CKD is not a direct result of CKD itself unless ESRD has been reached. There is evidence that mortality is increased in donors when large numbers are followed out over 20+ years. Mortality was higher in donors compared to healthy controls in a Norwegian cohort despite the fact that very few developed ESRD (Mjøen KI 2014). Even in the Ibrahim study (JASN 2016) in US white males and females an eGFR <60 ml/min was associated with a high relative risks of death even though few had developed ESRD. Short-term studies will miss this effect. In fact our results show that it is hard to detect any survival difference unless follow up is >20 years. Low eGFR is associated with increased mortality in all populations regardless of the cause. We assumed that the increased risk of mortality associated with CKD states was the same in donors and non-donors. The only exception was in the sensitivity analysis for isolated CKD (CKD not associated with proteinuria and diabetes mellitus). For CKD associated with diabetes and proteinuria we assumed that mortality in donors and non-donors would

be increased as they are in the general population. Rates of progression to proteinuria or to diabetes mellitus were not changed in donors and non-donors in this sensitivity analysis.
Minor comments
Methods: You do not state the cycle length of your Markov model.
Response: 1 year. Now added to paper.
Pg 10: “The mechanism by which low glomerular filtration rate CKD causes an increase in cardiovascular and all-cause mortality rate is not completely known” – is it known to be causal, or is it just a known association?
Answer: There are plausible (non-traditional risk factors, such has calcium/phosp disorders, anemia, etc) mechanisms, the increase appears to be independent of usual risk factors (DM, Hypertension, proteinuria, smoking etc), but to date most appear to be an association.
Pg 11: “black individuals appear to transition from CKD to ESRD at a faster rate than white individuals”; isn’t this dictated by the input parameters?
Answer: Correct.
Pg 12: the statement that “this would not be consistent with higher rates of ESRD in black males compared to white males in the general population” appears to contradict the estimates given in the previous sentence (that the added risk of ESRD was in 1 in 28 for white males compared to 1 in 22 for black males).
Response: One would have expected that the risk in white men would be 1/50 to 1/75 (rather than 1/26) if it was 1/22 for black men. This was clarified, thank you.

VERSION 2 – REVIEW

REVIEWER	Geir Mjøen Oslo University Hospital Norway
REVIEW RETURNED	13-Jun-2017

GENERAL COMMENTS	Please add a sentence or two at the end of the paper regarding the consequences of your findings for donor follow-up.

REVIEWER	Hassan Ibrahim Houston Methodist Hospital
REVIEW RETURNED	23-Jun-2017

GENERAL COMMENTS	I have no further comments

REVIEWER	Lois Kim Cambridge Centre for Health Services Research, Department of Public Health and Primary Care, University of Cambridge. UK.
REVIEW RETURNED	26-Jun-2017

GENERAL COMMENTS	Overall, although the authors have generally provided adequate responses to my initial comments in their letter, I do not feel that they have gone far enough in addressing these issues in the manuscript. My original comments are in italics below; my follow-up comments following the authors’ responses are added below in non-italics. Major comments 1. As with any disease progression model, the input parameters (along with model structure) are essentially what determine the

	outputs and results. It is therefore imperative that these are well-described both in terms of their source (and whether this is of sufficient quality, including any limitations regarding generalisability) and the form that they take. In this model, it is important to provide details about whether each transition probability differs by race/sex and whether they vary over time, for example, but this was not always clear. It would also be good to report the transition rates used, rather than just the relative increase in rates for donors as given in the appendix. Overall I did not feel there was sufficient detail to allow me to get a good feel for how the model results would be influenced by assumptions relating to the model inputs, which is key for interpreting the outputs. In looking at the key model results, I wanted to understand what was driving the higher loss of life-years in black donors for example, but this is difficult without knowing which transition parameters depend on race. I would also like to see more focus in the discussion from this perspective, looking at what aspects of the model are driving the between group differences. The authors provide some useful insight into the drivers of the difference in outcomes by race; however, no additions have been made to the manuscript to help the reader gain some understanding of this. This does not need to take the form of a full listing of all parameter estimates; some discussion about this would be sufficient. 2. There was also a lack of clarity about how the model was calibrated. It appears that the model was calibrated on the 15-year sex/race specific ESRD cumulative incidence. However, it was unclear how the various transition probabilities that feed into the ESRD state (e.g. the transition from diabetes to ESRD, the transition from CKD to ESRD) were estimated/calibrated to produce this. Furthermore, although the model may validate well on the 15-year incidence, there is no discussion of whether there was good validation over the whole 15 year period. Although the authors provide some detail about the calibration process in their reply, nothing further has been added to the manuscript about this. Furthermore, they also state that "to examine this uncertainty we also changed the multiplier in donors to match 15 year ESRD the upper and lower bounds of the 95% confidence interval and show that the outputs change modestly (see Table 1)." Whilst this does seem like a reasonable start to looking at uncertainty, I cannot see these results in Table 1 (I expected to find point estimates for remaining life-year estimates corresponding to different input estimates for the 15-year ESRD, but this is not given in Table 1). 3. It is clearly a simplification to treat CKD as one combined state, but there is no justification or discussion relating to this modelling choice. CKD would be better modelled as the constituent stages (3a to 5), which would each have substantially different quality adjustment and transition rates to other states, including mortality. I am glad to see that the authors have now acknowledged this as a limitation in the discussion section. I think it would also be valuable to include some consideration of the potential impact of this simplification on results. 4. It is not clear why Fig 1 does not include a state for hypertension and diabetes, as this is an extremely common combination of comorbidities. The authors reply "We agree and we assumed that patients with late onset diabetes would have hypertension.". However, I cannot see that this is stated in the paper (or appendix). 5. The investigation of uncertainty does not appear to include any

	propagation of uncertainty in parameter estimates through to parameter outputs. For example, the estimates of the relative difference in transition rates or quality of life adjustments will be estimated with uncertainty; this combined (joint) uncertainty across all the input parameters can then be propagated probabilistically through the model in order to provide uncertainty intervals about the results. There are some confidence intervals presented in Table 1, but it is unclear how these have been derived. The authors state in their response: “The uncertainty in the model was derived from Grams et al and is the 95% confidence interval for the cumulative ESRD event for their actual study”; however, I do not understand how the confidence interval for the primary outcome of the model, which is estimated by running the population through the particular model described in this paper (and which incorporates a range of different states and input parameters), can be estimated from another study? The CIs presented in Table 1 are extremely tight, suggesting a high level of certainty about the results; however, this is misleading since the real uncertainty about most of the input parameters appears to have not been incorporated at all (from the authors’ response: “At this point we do not have access to reasonable 95% confidence intervals and the distribution for many of the variables in actual donors although they exist for the general population from which the donors are derived”). The authors also state: “We added in the supplement loss of life years in an otherwise 40 year old white male ideal donors considering the uncertainty in the point estimated for ESRD progression associated with obesity and smoking. We hope these one sensitivity analyses help the readers and reviewer understand the magnitude of the uncertainly in these estimates”. However, these results are not sensitivity analyses that explore the impact of input parameter uncertainty – rather, they report results for different sub-groups of the population using different input parameter estimates. At the very least, there should be a discussion about the fact that uncertainty has not been addressed here. 6. Given that the mortality associated with CKD is generally not a direct result of the CKD itself (unless ESRD has been reached), it is an important assumption that the increased mortality for the CKD state is the same in the donor population as the non-donor population. This is acknowledged as a sensitivity analysis in terms of mortality, but it is not clear whether other aspects of this were also explored (progression to diabetes or proteinuria, for example). Related to this, the statements in the discussion (top of Pg 10) – “a significant percent of the loss of life was associated with CKD not ESRD” and “deaths during the CKD health state accounted for most of the projected increase in mortality and reduction in QALYs” need some further discussion. My point here is not that they don’t have increased mortality; the point is that this increase in mortality is due to comorbidities rather than CKD, i.e. the donors may not be expected to have the same level of comorbidities, and hence for a given CKD may have lower mortality than non-donors. Minor comments Pg 10: “The mechanism by which low glomerular filtration rate CKD causes an increase in cardiovascular and all-cause mortality rate is not completely known” – is it known to be causal, or is it just a known association? The authors acknowledge in their response that the majority of evidence to date is that this is an association rather than a causal relationship. However, there is no change to the text in the manuscript, which implies the relationship is causal.

	Pg 11: “black individuals appear to transition from CKD to ESRD at a faster rate than white individuals”; isn’t this dictated by the input parameters? The authors reply that my interpretation is correct, but have made no adjustment to the manuscript. I feel this reads as though it is a result, but this is not the case. The only reason black individuals transition from CKD to ESRD at a faster rate is because the authors have set this to be so in deciding the input parameter for this.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name
Geir Mjøen

Please add a sentence or two at the end of the paper regarding the consequences of your findings for donor follow-up.

Response: We have added 2 more sentences. Acknowledging that many these risk factors develop years after donation, short term follow up of kidney donors may be inadequate. Given that obesity and smoking are risk factors that might develop after donation, encouraging a healthy life style at donor evaluation and post donation is also important.

Reviewer 4.
1. The authors provide some useful insight into the drivers of the difference in outcomes by race; however, no additions have been made to the manuscript to help the reader gain some understanding of this. This does not need to take the form of a full listing of all parameter estimates; some discussion about this would be sufficient.

Response: The reasons why blacks are at higher baseline risks of ESRD, lower overall life expectancy and greater risk of ESRD with donation is a subject of much research (plausible explanations include social determinants of health, genetics, etc). The model is constructed to reflect this reality not to explain the underlying cause.

2. Although the authors provide some detail about the calibration process in their reply, nothing further has been added to the manuscript about this. Furthermore, they also state that “to examine this uncertainty we also changed the multiplier in donors to match 15 year ESRD the upper and lower bounds of the 95% confidence interval and show that the outputs change modestly (see Table 1).” Whilst this does seem like a reasonable start to looking at uncertainty, I cannot see these results in Table 1 (I expected to find point estimates for remaining life-year estimates corresponding to different input estimates for the 15-year.

Response: We think we now understand the confusion. The values in parentheses in Table 1 and 2 are not 95% CI per se but in fact represent the point estimates corresponding to different input estimates for the 15 year cumulative risk of ESRD in Table 1 and for the lifetime risks of ESRD in Table 2. We added a footnote to the Tables 1 and 2 noting this, ‘Values in parentheses represent the impact of higher and lower cumulative risks of ESRD taken from the upper and lower bounds of the 95% confidence intervals for lifetime cumulative risk of ESRD. (Supplement reference 11)’.

‘To evaluate a more conservative and more liberal estimate of remaining life years, lost life years post donation, remaining QALYs and lost QALYS post donation, we used higher and lower transition rates from Normal to CKD states that correspond to the upper and lower bound of the 95% confidence

interval of the projected cumulative risk of ESRD from study by Grams et al. (11)The range of outputs look tight because the risk of ESRD is small, occurs late and really has a very look tight because

3. I am glad to see that the authors have now acknowledged this as a limitation in the discussion section. I think it would also be valuable to include some consideration of the potential impact of this simplification on results.

Response: We have added the sentence,' . Since donors eventually have a greater risk of entering into a more advanced CKD state, this analysis may have underestimated the overall net loss of life years from nephrectomy' into the discussion.

4. The authors reply "We agree and we assumed that patients with late onset diabetes would have hypertension.". However, I cannot see that this is stated in the paper (or appendix).

Response: It is now in the supplement. Since the majority of patients with low GFR states and late onset DM have hypertension we assumed this in the model. (United States Renal Data System. 2016 USRDS annual data report: Epidemiology of kidney disease in the United States. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2016. Colosia AD, Palencia R, Khan S. Prevalence of hypertension and obesity in patients with type 2 diabetes mellitus in observational studies: a systematic literature review. Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy. 2013;6:327-338).

5. At the very least, there should be a discussion about the fact that uncertainty has not been

Response: The values in the Tables 1 and 2 in parentheses are not 95% CI as explained in #2. We added to the discussion, There are many variables and transition rates included in this model and addressing uncertainty in each or combinations of variables would require a much longer paper. The key uncertainties explored were risk of developing ESRD and the increased mortality associated with CKD states. Lower risks of developing diabetes mellitus and proteinuria, were also explored in ideal donors. But also included in methods that, 'We assumed that many future risks that can impact on life expectancy and ESRD such as cancer, obesity, smoking etc were not influenced by the act of kidney donation.

6. My point here is not that they don't have increased mortality; the point is that this increase in mortality is due to comorbidities rather than CKD, i.e. the donors may not be expected to have the same level of comorbidities, and hence for a given CKD may have lower mortalitythan non-donors.

Response: We disagree. A low GFR independent of co-morbidity is associated with an increase in CVD death has been shown in many studies.

Minor.
• The authors acknowledge in their response that the majority of evidence to date is that this is an association rather than a causal relationship. However, there is no change to the text in the manuscript, which implies the relationship is causal.
Response : We changed the sentence in question to 'CKD is associated with' rather than assume a cause.

• The only reason black individuals transition from CKD to ESRD at a faster rate is because the authors have set this to be so in deciding the input parameter for this.

Response: The reason for the inputs for blacks are higher is that by not doing so the outputs would have blacks and whites having the same risk of ESRD which is not true. We have deleted the

sentence disturbing to the reviewer. This reflects the time spent with CKD relative to ESRD. Black individuals appear to transition from CKD to ESRD at a faster rate than white individuals.